The Effect of Endoplasmic Reticulum and Reductive Stress on the Human Dermal Fibroblast Proteome

CARNE, NAOMI,ANGHARAD

The Effect of Endoplasmic Reticulum and Reductive Stress on the Human Dermal Fibroblast Proteome

Naomi Angharad Carne

2018

A thesis submitted for the degree of Doctor of Philosophy

Department of Biosciences
Durham University

Abstract

Dermal fibroblasts are responsible for the secretion of extracellular matrix (ECM) components that support the structural integrity of the skin. Alterations to the ECM have been implicated in many skin diseases including systemic sclerosis and fibrotic disorders, as well as wrinkle formation and wound healing in the aged phenotype. The endoplasmic reticulum (ER) is responsible for the production and quality control of secreted proteins, and perturbations to its correct function could therefore lead to aberrant ECM deposition from dermal fibroblasts. ER stress occurs when homeostasis of this organelle is imbalanced, which can be prompted by the effects of redox agents that disrupt the careful redox balance within the ER lumen. Previous research has often focused on the effects of oxidising agents that lead to oxidative stress within the cells, however little is known about the effects of reductants (and therefore reductive stress). Reductants are present in pollutants, depilatory creams and some cosmetics yet relatively little is known about their potential effects on the skin. This thesis aims to investigate the effect of reductive stress on dermal fibroblasts, looking first at signalling responses and then investigating changes that occur at the proteomic level. In the final chapter a comparison is made between the proteomic response to reductive stress by DTT and oxidative stress by UV-A radiation. The implications of these findings are discussed in the context of fibroblast functions in the skin.

Table of Contents

Table of Figures

Chapter 1

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Table of Tables

Chapter 2

Chapter 3

Chapter 5

Chapter 6

Table of Abbreviations

Specific proteins of interest are referred to using their gene name throughout this thesis and are not listed here. Amino acids are referred to using the three-letter code.

2D DiGE	2D Difference Gel Electrophoresis
2D PAGE	2D Polyacrylamide Gel Electrophoresis
ABC	Ammonium bicarbonate
APCI	Atmospheric Pressure Chemical Ionisation
b-FGF	basic Fibroblast Growth Factor
BM	Basement Membrane
BSA	Bovine Serum Albumin
C_{12}FDG	5-dodecanoylaminofluorescein di-β-D-galactopyranoside
CE	Cornified Envelope
cmc	Critical micelle value
CV	Coefficient of Variance
DDA	Data-Dependant Acquisition
DIA	Data-Independent Acquisition
DTT	Dithiothreitol
DWAT	Dermal White Adipose Tissue
E	Embryonic day
ECM	Extracellular Matrix
EDC	Epidermal Differentiation Complex
EGF	Epidermal Growth Factor
eIF	Eukaryotic Initiation Factor
EMT	Epithelial to Mesenchymal Transition
ER	Endoplasmic Reticulum
ERAD	ER Associated Degradation
ERO	Endoplasmic Reticulum Oxidoreductase
ESI	Electrospray Ionisation
FASP	Filter Aided Sample Preparation
FCS	Foetal Calf Serum
FDR	False Discovery Rate

FKBP	FK Binding Protein
FT-MS	Fourier Transform Ion Cyclotron Mass Spectrometry
GAG	Glycosaminoglycan
GO	Gene Ontology
Gpx	Glutathione peroxidase
H_2O_2	Hydrogen peroxide
HD	Hemidesmosome
HGPS	Hutchinson Gilford Progeria Syndrome
HLCC	Hydroxylysine-aldehyde Collagen Crosslinks
HMGR	3-hydroxy-3-methylglutaryl acetyl-coenzyme-A reductase
HSPG	Heparin Sulfate Proteoglycans
IEF	Isoelectric Focussing
iTRAQ	Isobaric Tags for Relative and Absolute Quantification
LC	Liquid Chromatography
LCC	Lysine-aldehyde Collagen Crosslinks
LC-MS/MS	Liquid Chromatography - tandem Mass Spectrometry
LMNG	Lauryl Maltose Neopentyl Glycol
m/z	Mass to charge ratio
MALDI	Matrix Assisted Laser Desorption/Ionisation
MEM	Minimum Essential Media
MHC	Major Histocompatibility Complex
MMP	Matrix Metalloproteinase
MS	Mass Spectrometry
NAC	N-acetylcysteine
nHDF	Normal Human Dermal Fibroblasts
OST	Oligosaccharide transferase
P-	Phosphorylated
P&G	Procter & Gamble
PBS	Phosphate Buffered Saline
PC	Principal Component
PCA	Principal Component Analysis
PDGF	Platelet Derived Growth Factor
PDGFR	Platelet Derived Growth Factor Receptor

PDI	Protein Disulfide Isomerase
PPIase	Peptidyl-Prolyl *cis-trans* Isomerase
RIDD	Regulated IRE1-Dependant Decay
RIPA	Radioimmunoprecipitation Assay
ROS	Reactive Oxygen Species
RTK	Receptor Tyrosine Kinase
RT-PCR	Reverse Transcription - Polymerase Chain Reaction
SAβGal	Senescence-Associated β-Galactosidase
SASP	Senescence-Associated Secretory Phenotype
SCX	Strong Cation Exchange
SMC	Structural Maintenance of Chromosomes
SRM	Selected Reaction Monitoring
SRP	Signal Recognition Particle
SSc	Systemic Sclerosis
SS	Splice Site
SWAT	Subcutaneous White Adipose Tissue
SWATH	Sequential Window Acquisition of all Theoretical Fragment Ion Spectra
TAC	Transit Amplifying Cell
TBS	Tris Buffered Saline
TBS-T	Tris Buffered Saline with Tween
TIMP	Tissue Inhibitor of Metalloproteinase
TMPA	Terminally Misfolded Protein Aggregates
TNFα	Tumour Necrosis Factor alpha
TOF	Time of Flight
TX100	Triton X-100
UPR	Unfolded Protein Response
UV	Ultraviolet

Acknowledgments

First, I must thank my supervisor Adam Benham for the opportunity to complete this PhD, and for his help and guidance throughout the past four years. I must also thank Adrian Brown for all his work in the proteomics facility to support this thesis. Thanks also go to Mike Flagler and the team at P&G for the generous funding and their input throughout the project. I would also like to thank my thesis committee for their input and advice, and especially Heather Knight for being a friendly face and source of encouragement when I needed it most.

Secondly, my thanks go to everyone in Lab 8 for your friendship and help over the years, and for putting up with my ridiculous time-courses in the culture hood. Thanks especially to Steven for all your help with anything microscope related – good luck with your PhD, I look forward to seeing your results. Sarah – your understanding and friendship, especially since your move from Stockton, has been invaluable, I wish you all the best with the writing up and I promise it's not as bad as you think.

Next, I need to thank my friends outside the lab, who have kept me going all this time. I can't list them all, but thanks especially go to Beano for always being a source of comfort and optimism, to Vicki for her help and encouragement throughout, and to Rachel for all her support and Monday night phone calls. Also, to my friends at the Cobwebs Orchestra in Spennymoor for cheering me up on a Wednesday evening, and to St Giles for being a home away from home whenever I needed it.

Special thanks go to James for always having faith in me, encouraging me and for comforting me when the going got tough. You never let me give up and I am eternally grateful. Also, thanks for putting up with the East coast mainline for so long!

Finally, I would like to thank my family. Mum, Dad and Eleanor for your unwavering love, support and encouragement: I couldn't have done it without you and I hope I've done you proud. Yn olaf ond nid lleiaf, diolch yn fawr hefyd i fy mamgu am ei chariad a'i chefnogaeth bob amser.

Chapter 1: General Introduction

1 Introduction

This chapter serves as an introduction to the three main themes of this thesis. First, it will describe the structure of the skin, and the function of each of the three main layers of this tissue: epidermis, dermis and hypodermis. This will serve to contextualise the role of the dermal fibroblast, the focus of this study. Secondly, the function of the endoplasmic reticulum (ER) and the consequences of ER stress will be described with reference to its role in dermal fibroblasts. Finally, an introduction to proteomics and the importance of this technique in providing a hypothesis-independent understanding of cellular responses to stress will be given.

1.1 Structure and function of human skin

The skin is widely regarded as the largest organ of the human body, and functions at the interface with the external environment. Indeed, every organism on the planet can be thought of as having a 'skin': a membrane that separates their contents from the outside world, specialised and well adapted to each specific environment.

Human skin functions to maintain homeostasis both with respect to alterations in the external environment, as well as retention of sufficient water that could otherwise be lost through evaporation (trans-epidermal water loss). Nerve endings within the tissue provide information about touch and temperature, which are transmitted to the brain to stimulate appropriate responses. The skin must also provide protection against physical (i.e. mechanical injury or ultraviolet (UV) radiation), chemical (i.e. irritants or allergens) and microbial (i.e. bacterial, fungal or viral) assaults. Incorrect formation of the skin barrier is associated with several human diseases including, for example, eczema, psoriasis and the rare but severe ichthyosis disorders characterised by the dry, "fish-scale" appearance of skin[1].

The skin is divided in to three layers: the hypodermis at the bottom, immediately above muscular tissue; the dermis in the middle and the epidermis at the top. The dermis and epidermis are separated by the basement membrane (BM). In addition to the cells within these layers, various adnexal structures exist including the hair follicle, sweat and sebaceous glands (Figure 1.1). The development of a complete structure, able to maintain homeostasis and prevent excessive water loss, occurs at around 34 weeks during embryo development in humans. Therefore, babies born before this date require careful incubation until the skin development is complete[2].

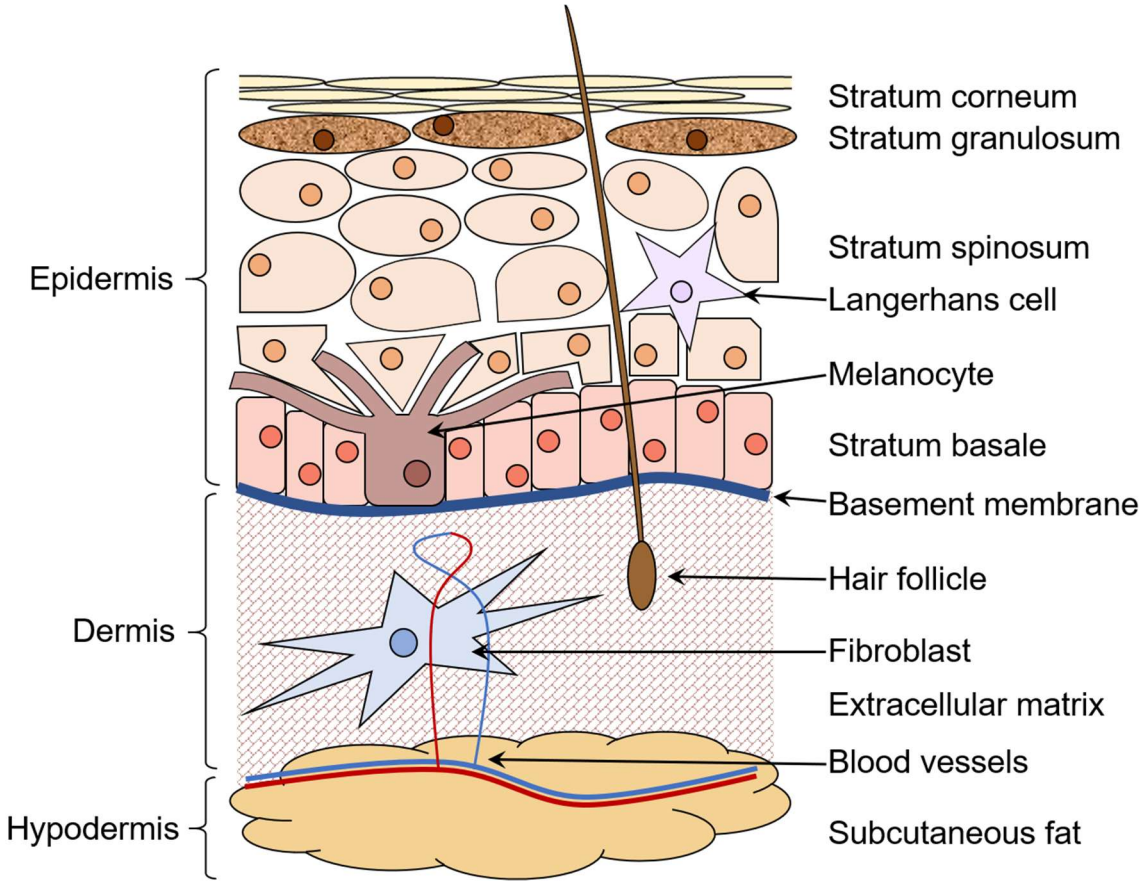


Figure 1.1: Structure of the skin. Diagram depicting the many layers of human skin. The upmost layer, the epidermis, is made of several layers of differentiating keratinocytes that become flatter and cornified as they migrate away from the basement membrane. The dermis lies beneath the basement membrane and is the location of the fibroblast cells. The hypodermis is the deepest layer of the skin and is associated with the subcutaneous fat. Self-drawn adaptation from: Visscher, M. & Narendran, V. Neonatal Infant Skin: Development, Structure and Function. Newborn and Infant Nursing Reviews 14, 135–141 (2014).

1.1.1 Epidermis

The main cell-type found in the epidermis is the keratinocyte. These are highly replicative cells that go through a process of terminal differentiation as they age. New keratinocytes are generated in the deep epidermis, next to the BM and, as differentiation and ageing of these cells occurs, they progress to the outermost layer eventually being shed as dead skin.

Keratinocytes are characterised by their production of keratins, the major intermediate filament protein, which can also be used to define the stage of differentiation of each cell in the epidermis. As such, they have become popular as biomarkers of skin tumours where they can be used to define the exact epidermal location from which the cancer arose, as the tumour cells maintain the expression patterns of their origin[3]. Keratins are comprised of an α-helical rod domain, with a head and tail domain on either side, and are classified into two groups according to their isoelectric point: type I (acidic) and type II (basic). The α-helical domain of two keratins, one of each type, come together into a coiled-coil formation of the intermediate filament. As such, keratins are usually found in pairs. In the basal layer of the epithelium keratins K5 (type II) and K14 (type I) are the most strongly expressed and are fixed to desmosomes and hemidesmosomes that attach cells to the BM[3]. During differentiation, keratin expression switches to the production of suprabasal epidermal keratins K1 (type II) and K10 (type I). This pair of keratins form particularly dense bundles which are associated with the keratinocytes seen in the upper layers of the epidermis[4].

The process of terminal differentiation has been studied *in vitro* using an experimental model where keratinocytes are grown in culture, under conditions in which they form stratified sheets. In these sheets the basal layer is attached to the culture substrate and additional layers form on top of this, undergoing terminal differentiation before being shed into the culture media[5]. A reduction in cell-substratum contact, as would be the case in a crowded layer of cells on the BM, is thought to trigger the process of terminal differentiation within keratinocytes[5]. The differentiation process is also mediated by a strict gradient of calcium, which increases with distance away from the BM. Restriction of calcium availability prevents keratinocyte differentiation into a stratified layer *in vitro*[6].

Terminally differentiated cells are replaced by proliferation of a population of stem cells, which may replicate in an unlimited fashion, and their daughters, termed transit-amplifying cells (TAC), which have a more limited replicative capacity but are pre-destined for terminal differentiation[7]. These cells are large and columnar, forming the stratum basale. The two cell types are distinguished by their surface expression of β1-integrin: true keratinocyte stem cells express much higher levels of β1-integrin than TACs[8]. Thus, the distribution of stem cells in the epidermis has been evaluated to reveal clusters of stem cells, located at specific sites of the epidermal-dermal junction, and these are connected by a network of TACs[9].

Amongst the undifferentiated keratinocytes of the stratum basale are melanocytes. Melanocytes are defined by the presence of a unique organelle, the melanosome, which produces the pigment melanin that confers photoprotection against UV radiation to the skin[10,11]. Melanosomes originate from the trans-Golgi network and migrate towards the cell periphery as they mature, eventually being transferred to surrounding keratinocytes (although the precise mechanism of this transfer is still under debate)[12–14]. It is melanin, derived from melanosomes produced by the melanocytes that is responsible for the pigmentation of skin known as tanning.

As the columnar epithelial cells detach from the BM and migrate upwards through the epithelium, there is a gradual change in shape such that in the stratum corneum the cells are completely flattened and densely packed. The differing layers within the epithelium can be defined by the types of keratin proteins expressed, and the nature of cell-cell junctions found. The stratum spinosum lies above the stratum basale and consists of polyhedral cells, punctuated by protruding desmosomes that give cells their characteristic "spiny" appearance. It is in this layer that lipid containing lamellar bodies first start to develop[15]. Also within this layer are the Langerhans cells, which (in healthy epidermal tissue) are the major antigen-presenting cell[16].

Langerhans cells are identified by the expression of a C-type lectin receptor named 'langerin' on their cell surface[16,17]. They have long been thought of as antigen-presenting cells that, upon receipt of danger signals, mature and migrate to the lymph nodes and present antigens to naïve resting T-cells and lead to stimulation of an immune response[16]. Since the millennium, new studies that took advantage of genetic engineering technologies in mice models have muddied the waters with respect to the specific function of Langerhans cells. Allan et al., showed that in cutaneous herpes virus infections Langerhans cells were not responsible for the presentation of antigen to T-cells[18]. In addition, Kaplan et al., demonstrated that mice that were depleted of Langerhans cells in fact showed enhanced responses to a model of contact dermatitis previously thought to involve Langerhans cells[19]. Kautz-Neu et al., have since implicated Langerhans cells in the dampening of the immune response to *Leishmania major* infection[20] suggesting that Langerhans cells may play a more complex role in epidermal immune responses than simply in the activation thereof (reviewed by Romani et al.[16]).

As the cells continue their migration through the epidermis, they undergo a final flattening and loss of organelles in the granular layer (stratum granulosum). Cells in this layer are characterised by the presence of many keratohyalin granules the main component of which is the giant filaggrin pre-cursor: pro-filaggrin. As the keratinocytes undergo the final stages of terminal differentiation, pro-filaggrin is cleaved into filaggrin monomers that are released from these granules and associate with keratin intermediate filament proteins to form tightly packed keratin macrofibrils[21]. These

facilitate the flattening of the cells and are crosslinked to form a keratin mesh onto which the cornified envelope that characterises the stratum corneum is attached[21,22].

Lamellar bodies, also found within cells of the granular layer, contain folded sheets of lipid layers that are secreted by exocytosis during differentiation at the interface between the stratum granulosum and the stratum corneum. Phospholipids, stored within the lamellar bodies of the cells, are released into the extracellular space and the cell membrane alters to become rich in ceramide lipids, cholesterol and fatty acids[23]. These sheets of hydrophobic phospholipids act as a waterproofing barrier that protects the skin from excessive water loss. In addition to lipids, the lamellar bodies also release a range of proteins and enzymes into the extracellular space. These include proteases important for the correct sloughing of corneocytes from the skin surface as well as anti-microbial peptides that enhance the skins protection against pathogen attack[24].

Eventually, the stratum granulosum gives way to the stratum corneum in the upmost layer of the epidermis. This is a layer of anuclear cells, which are rich in keratin and surrounded by the cornified envelope (CE), immediately below the ceramide lipid-rich layer of the modified cell membrane (termed the lipid envelope). The CE is a layer of cross-linked proteins, most of which are products of genes in the epidermal differentiation complex (EDC) on chromosome 1, that is insoluble in both detergent and reducing agent[25]. CE formation is initiated by the expression of envoplakin and involucrin which associate around the cell membrane and are crosslinked by transglutaminase enzyme 1 (TGase 1) to form the scaffold[26]. This scaffold is then re-enforced by the recruitment of other CE proteins including loricrin and small proline-rich proteins that are crosslinked by transglutaminase enzyme 3 and linked to the scaffold by TGase 1[26]. Keratin macrofilaments are also crosslinked to this structure, reinforcing its strength and rigidity[26].

Once terminally differentiated keratinocytes, now corneocytes, reach the surface of the skin they are shed and lost to the environment. The shedding of corneocytes as dead skin is regulated by the breakage of desmosome junctions between cells. In the desmosomes, intracellular intermediate filaments are connected to the intercellular desmosomal cadherins by desmoplakin and plakophilin proteins[27]. These are rigid structures that provide strength and stability to the stratum corneum that can resist mechanical stress. Perturbations to the desmosome structure or formation is implicated in many diseases including, for example, epidermolysis bullosa simplex characterised by excessive skin blistering[27]. Corneodesmosomes exist between corneocytes while transition desmosomes exist between corneocytes and the cells of the stratum granulosum. Proteolytic cleavage of these junctions regulates when and where cells are lost from the skin surface, the proteases being kept inactivated when not required by inhibitors such as LEKT-1. LEKT-1 is a serine protease inhibitor implicated in the regulation of shedding the stratum corneum. The importance of this regulation is demonstrated by diseases such as Netherton syndrome, where mutations in

LEKT-1 result in the premature cleavage of desmosomal junctions leading to the loss of the stratum corneum, skin inflammation, allergy and scaling[28,29].

1.1.2 Basement membrane

The BM is a specialised ECM that forms on the surface of competent epithelial cells and separates the epidermis from the dermis. It protects tissue from physical stresses and provides a way to mediate the communication of distant signals between compartments. The BM is primarily composed of laminins, nidogens, heparin sulfate proteoglycans (HSPG) and type IV collagens. Laminins associate with sulfated glycolipids, integrins and dystroglycan on the cell membrane of competent cells and this initiates the self-assembly of the BM structure[30,31]. Nidogens bind to the coiled coil domains of laminin proteins, and to type IV collagens to form a stabilising bridge that attaches collagen to the laminin network. The collagen proteins polymerise to form a network above the laminins and HSPGs bind to the mesh to form additional contacts with cell surface receptors[31]. Loss of BM components is lethal and prevents complete and correct development of the embryo[32].

Laminins exist as heterotrimers of an α, β, and γ chain with each consisting of several globular and rod-like domains. Heterotrimer assembly occurs within the cell, although some laminins are subject to further proteolytic processing after secretions into the ECM[33]. Laminin epidermal growth factor-like domains, which bind to nidogen-1 and -2, are found in γ1 and γ3 chains and these are thought to be essential for BM assembly. Nidogens are not thought to be essential to the overall BM architecture and assembly, rather they act as stabilising molecules and their breakdown is implicated as a key event in BM degradation[34].

Historically, type IV collagens have been thought to play a solely structural role in the BM, but in the last 18 years evidence of a more active role has become apparent. A role in cell signalling has been defined as type IV collagens were shown to activate the G-protein coupled receptor GPR126, which is essential for organ development[35]. In melanoma cells, the level of glycosylation on type IV collagen chains has been shown to modulate integrin binding and therefore cell adhesion[36]. Promotion of type IV collagen-integrin interactions via the globular NC1 domain at the end of collagen triple-helical chains has also been shown to inhibit angiogenesis and tumour growth *in vivo* suggesting a role for these collagens in the inhibition of tumour growth[37]. This is supported by investigations that demonstrate a loss of type IV collagens in the early stages of cancer invasion and remodelling of type IV collagen chains in many cancer cell types. Recently, this has been associated with the hypermethylation, and therefore silencing, of the promoter region for two of the type IV collagens α5(IV) and α6(IV)[38] implicating a loss of type IV collagen in the tumour progression and metastasis.

Epithelial cells are bound to the BM through hemidesmosome (HD) junctions that link the keratin intermediate filaments of cells to the BM by integrin-mediated attachments[39]. In stratified epithelia, such as that of the skin, the proteins involved in HD junctions are integrin α6β4, plectin isoform 1a, tetraspanin, bullous pemphigoid antigen (BPAG) 1e and type XVII collagen. K5/K14 intermediate filaments bind to plectin and BPAG1e within the cell, linking to the transmembrane complexes of integrin $\alpha 6\beta 4$, type XVII collagen and tetraspanin that bind laminins on the extracellular side of the plasma membrane[39]. The laminins find anchorage fibrils of type IV collagen to create a firm connection between cell and BM. Mutation in any of the components of the HD junction are implicated in blistering diseases of the epidermolysis bullosa family (see section 1.1.1).

1.1.3 Dermis

The function of the epidermis is supported by the cells and ECM structures of the dermis that lies beneath. The dermis is divided into two sections: papillary, which is immediately beneath the BM, and reticular that lies deeper in the skin, above the hypodermis. These are separated by a vascular plexus termed the rete subpapillare: a structure of blood vessels that supply the cells of the skin with the nutrients required to survive. As the epidermis is void of vascular tissue, it obtains the essential nutrients from the blood through diffusion from capillaries in the dermis. Indentations into the epidermis called dermal papillae extend the surface area for interactions between epithelia and dermis, and provide routes into which capillaries can extend, increasing the availability of nutrients (Figure 1.2)[40]. The rete subpapillare that demarcates the boundary between papillary and reticular dermis is fed by a further vascular plexus, the rete cutaneum, at the boundary between the dermis and hypodermis[40].

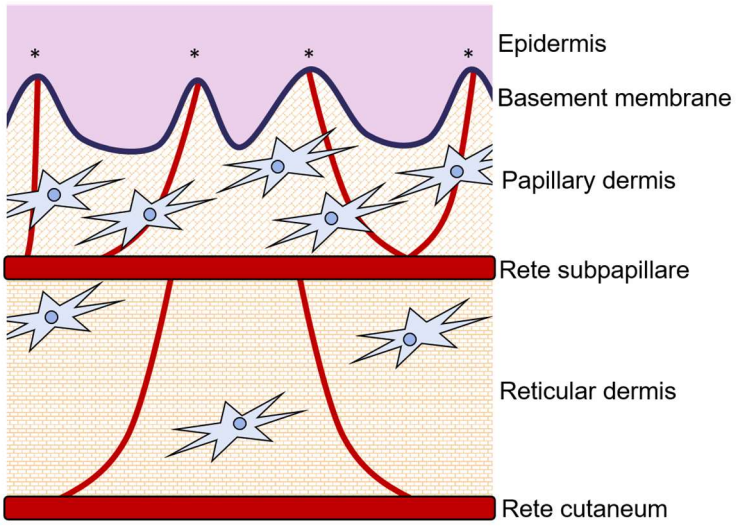


Figure 1.2: Structure of the dermis. Diagram depicting the different levels of the dermis. The basement membrane is folded into dermal papillae (marked by ∗) on top of which lies the epidermis. The dermis is divided into papillary and reticular layers by the rete subpapillare and rests on the rete cutaneum between the dermis and hypodermis at the bottom of the skin. Self-drawn adaptation from: Sorrell, J. M. & Caplan, A. I. Fibroblast heterogeneity: more than skin deep. J. Cell. Sci. **117,** 667–675 (2004).

The dermis is characterised by the expanse of organised ECM found within it that provides skin with the ability to resist both compressive and contractive forces, as well as providing tensile strength. The two main components of the ECM are glycosaminoglycans (GAG), which are often linked to proteins to form proteoglycans, and fibrous proteins that are either structural (e.g. collagen/elastin) or adhesive (e.g. fibronectin/laminin). The GAGs and proteoglycans form a hydrated gel that resists compressive forces, and the fibrous proteins that provide tensile strength are embedded within this[4].

GAGs are unbranched polysaccharide chains of repeating disaccharide units, one residue of which is always either N-acetylglucosamine or N-acetylgalactosamine. These polysaccharide chains are inflexible and carry many negative charges; as such they occupy a vast amount of space and attract many osmotically active cations thus drawing water in to create a hydrated gel. This gel not only provides turgor required to resist compression but also allows the rapid diffusion of water soluble molecules to provide nutrients and permits migration of cells within the ECM. With the exception of hyaluronan, all GAGs are covalently linked to protein in the proteoglycans. These GAGs can bind and modulate the effect of signalling molecules and other secreted proteins such as proteases. Syndecan is a specific transmembrane proteoglycan on the surface of fibroblasts that acts as a receptor for fibronectin and collagen to mediate cell attachment to the matrix, or migration through it. It may also bind signalling molecules such as fibroblast growth factor and present this to the cell, thus acting as a co-receptor with cell-surface receptor proteins[4].

The structural proteins collagen and elastin form an organised matrix within this gel to provide the ECM with tensile strength and stretching capacity. Collagens are synthesised as procollagen α-peptides, three of which come together to form the distinctive triple-helix. The triple helix is stabilised by hydroxylation of proline and lysine residues that occurs during protein maturation in the ER and permits formation of interchain hydrogen bonds. After secretion, these triple helical units mature and assemble into ordered polymers called collagen fibrils, which often further associate into bundles termed collagen fibres. The fibrillar collagens (type I, II, III, V and XI) are organised into matrices by a combination of fibrillar associated collagens (type IX and XII) and the contractile action of fibroblasts attached to the dispersed fibrils. The pressure exerted by fibroblasts on the matrix contracts collagen fibres into sheets and draws them into cables to increase the matrix density[4].

Elastic fibres formed from the protein elastin provide the skin with a stretching capacity that is moderated by the interaction with the collagen fibre network, which controls stretch and prevents tearing. Elastin is a hydrophobic protein that, like collagen, is rich in proline and glycine but is not glycosylated or hydroxylated. After secretion, elastin monomers become crosslinked to generate a

network of fibres and sheets that can be stretched when required to afford skin the flexibility required, for example to stretch over moving joints[4].

Cells are connected to the ECM by adhesive proteins such as fibronectin. Fibronectin is a large, dimeric glycoprotein made of several functionally distinct domains that bind different entities such as collagens and cells. Fibronectin filaments assemble on or near the surface of fibroblasts and interact with the intracellular cytoskeleton via integrins. Integrins are transmembrane heterodimers that act as cell adhesion proteins to attach cells to the ECM. When integrin binds its matrix ligand, assembly of an intracellular attachment complex is triggered that links the integrin to the cytoskeleton, most commonly to the actin filaments[4]. Different combinations of α and β chains come together to form integrins of differing specificities that will interact with specific ECM or cytoskeletal components.

Fibroblasts are the primary cell type found within the dermis, and can be categorised into three distinct subpopulations according to their location: papillary, where they are more densely packed; reticular where there are fewer cells amongst a more organised ECM; and those that are associated with the hair follicle[40]. They are responsible for the secretion of ECM components such as collagen and glycosaminoglycans and are therefore integral to maintaining the structure and elasticity of the skin[41]. They also play an important role in wound healing and skin repair, which has led to them becoming the focus of many studies, including this thesis, regarding the role they may play in fibrosis, skin ageing and repair.

Reticular and papillary fibroblasts demonstrate slightly different morphologies in culture: a squarer appearance is associated with the reticular fibroblasts while papillary fibroblasts are said to have a spindle shape[42]. It has been proposed that these unique subpopulations of fibroblasts help to organise the complex microvasculature that lies beneath the avascular epidermis[43]. Fibroblasts that are found in the papillary dermis are thought to provide a microenvironment that supports the microvasculature found at this location. The formation of a supportive tubule structure is only robustly seen with papillary dermis fibroblasts, while fibroblasts isolated specifically from the reticular dermis take longer to produce supportive tubular networks, which are also sparser, shorter and less branched[43]. In addition, the two subpopulations demonstrate different characteristics in culture: fibroblasts derived from the papillary region divide at faster rates than site-matched reticular fibroblasts, and attain a higher cell density as they are not fully contact inhibited[44,45].

More recently, new evidence suggests that papillary and reticular fibroblasts derive from separate genetic lineages, as well as being spatially separate[46]. Studies of gene expression patterns in mouse skin throughout development have identified several specific markers for the papillary and reticular subpopulations of fibroblasts[47] (Figure 1.3). Fibroblasts which differentiate to form the lower,

reticular, dermis are identified by expression of delta-like homolog 1 (Dlk1) and Sca1 after embryonic day (E) 16.5[47]. Fibroblasts which will form the upper, papillary dermis express B-lymphocyte-induced maturation protein 1 (BLIMP1), a transcriptional regulator of terminal differentiation[48], and CD26, a cell surface serine exopeptidase[49], after E16.5[47]. However a more recent study has suggested that the expression of CD26 is not necessarily specific to the fibroblast location in the dermis, but rather to its role in matrix deposition after injury[49]. Thus, individual fibroblast populations have specific roles as well as locations, which have not yet been fully elucidated. It is known, however, that reticular fibroblasts are responsible for the bulk of collagen fibril and ECM production within skin tissue, and it is these cells that are recruited initially in wound repair, providing an explanation for the ECM rich scar tissue that is observed after wound healing[46].

The reticular and papillary regions differ also in the content and organisation of their extracellular matrices. The papillary dermis ECM contains thin collagen fibres of type I and type III that are poorly organised, whereas the reticular dermis has much less type III collagen but thicker and well-organised bundles[40]. Decorin, a proteoglycan involved in matrix assembly, is more abundant in the papillary dermis whereas versican, a much larger chondroitin sulfate proteoglycan, is more abundant in the reticular dermis[40]. In addition, collagen type XII and type XVI, and tenascin-C are found in the papillary dermis but collagen type IV and tenascin-X are the characteristic types of these proteins found in the reticular dermis[40].

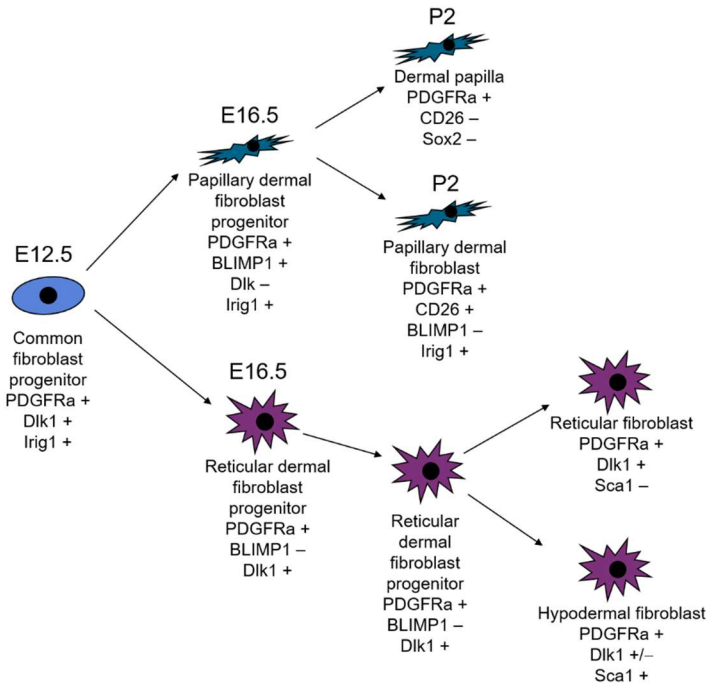


Figure 1.3: Biomarkers of different fibroblast lineages Diagram to depict the differentiation of fibroblasts from a common progenitor (**blue**) into upper, papillary (**green**) and lower, reticular (**purple**) lineages with specific genetic markers shown at each time point, either embryonic (E) or postnatal (P) days. Self-drawn adaptation from: Driskell, R. R. & Watt, F. M. Understanding fibroblast heterogeneity in the skin. Trends Cell Biol. 25, 92–99 (2015).

1.1.4 Hypodermis

The hypodermis is a layer of fat-storing adipocytes beneath the reticular dermis and rete cutaneum of the skin. The term 'hypodermis' has historically referred to all white adipose tissue below the dermis, but it is now clear that there is a distinction, based on adipocyte morphology and physiology, between dermal white adipose tissue (DWAT), the "true" hypodermis immediately below the reticular dermis, and subcutaneous white adipose tissue (SWAT) that lies beneath[50]. The adipocytes in the DWAT develop alongside fibroblast development, but distinct from the development of adipocytes in the SWAT[50]. This distinction had been previously demonstrated by Kelley et al., who showed a difference between these two layers with respect to their association with insulin resistance, where SWAT was strongly associated but DWAT was not[51]. Adipocytes are necessary for the recruitment of fibroblasts and reconstruction of the dermis following wound healing[52]. The DWAT cells are also involved in the hair follicle cycle, thermal regulation and protection again pathogen attack[53]. Thus, the hypodermis, or more accurately DWAT, has a specific and important role in maintaining skin function.

1.2 The Endoplasmic Reticulum and ER stress

1.2.1 Function of the ER

The ER is a major perinuclear organelle that consists of a network of membrane-bound tube structures called cisternae and may be decorated with ribosomes (rough ER) or not (smooth ER). The smooth ER is involved in the synthesis of lipids, phospholipids and steroids but is not usually abundant, except in specialised hormone secreting cells such as the Leydig cells of the testis for testosterone secretion[4]. Rather, in most cell types, the rough ER has regions of membrane that are void of ribosomes, often termed transitional ER. Rough ER synthesises most secreted, transmembrane and luminal-resident proteins. Separated from the cytosol, it provides a unique environment to support the correct folding and assembly of secretory proteins, particularly in the formation of disulfide bonds.

Translation of mRNA into proteins by ribosomes releases a nascent polypeptide chain that must correctly fold for the protein to assume its active role within the cell. That is, it must assume its native state. In 1973 Anfinsen et al., concluded that the native state of a protein was determined by the amino acid sequence of its peptide chain, that folded into the most thermodynamically stable conformation based on the surrounding solvents[54]. Levinthal noted that there were a vast array of possible conformations available to the nascent polypeptide and it was not possible that all of these were surveyed during a folding process that occurs, for the most part, in a microsecond time-frame[55]. Thus, selection of a native state could not be due to random or systematic sampling of all possible conformations, there must be additional factors that influence a proteins ability to achieve its native fold. It was later determined that the hydrophobicity of amino acids was a major driving force in the correct folding of proteins. Peptides that had the same pattern of hydrophobic and hydrophilic amino acids but ultimately differed in their absolute amino acid sequence obtained that same folding arrangement[56]. Hydrophobic residues tend to be buried in a central core of the protein, protected from aqueous solvent in the cytosol while hydrophilic residues are open to the environment. A range of other intermolecular interactions influence and stabilise the native structure: van der Waals forces, hydrogen bonds, disulfide bridges and ionic interactions all play their part[57]. Correct folding of proteins is also mediated by proteins such as the molecular chaperones (described in more detail below) which protect exposed hydrophobic sites in folding intermediates and aid protein folding by preventing aggregate formation[58].

Secreted and transmembrane proteins utilise a specific folding environment in the ER that promotes formation of disulfide bonds and provides a quality control mechanism to prevent secretion of terminally misfolded proteins. This is important as there are very few chaperones that exist in the extracellular space and as such there is little or no capability to repair misfolded proteins following secretion[59]. Protein folds in secreted proteins tend to have more disulfide bonds than

cytosolic proteins and therefore their correct folding is promoted by the ER environment. For example, the epidermal growth factor (EGF) – like fold is only found in secreted or transmembrane proteins and consists of two, two-stranded β sheets joined by a loop and contains 3 disulfide bonds[60,61]. A carefully maintained redox balance within the ER lumen permits both the oxidation and reduction that allows rearrangement of disulfide bonds to occur during protein folding[62,63] (see section 1.2.2.4).

1.2.2 Protein folding in the ER

Most proteins destined for the secretory pathway are identified during translation by a signal peptide at their N-terminus. The actual sequence of the signal peptide is highly variable although always consists of a series of hydrophobic residues, flanked on the N-terminal side by positively charged residues, and on the C-terminal side by polar residues that provide a cleavage site for the peptide's eventual removal[64,65]. As this sequence emerges from the ribosome it is recognised by the signal recognition particle (SRP) which delivers the ribosome, in complex with mRNA and emerging nascent chain, to an SRP receptor on the ER membrane (Figure 1.4). SRP binding also inhibits further elongation of the peptide emerging from the ribosome until it has been successfully docked to the SRP receptor[64]. This prevents translation of secretory pathway proteins in the wrong cellular compartment which might otherwise occur due to a rate-limiting number of SRP receptors on the ER membrane[66].

Following binding of SRP/ribosome/nascent peptide complex to the SRP receptor on the ER membrane, SRP is recycled having positioned the ribosome on top of the Sec61 translocon, allowing translation and elongation of the peptide to resume[64]. Sec61 is made up of three subunits α, β, and γ which assemble to provide an aqueous pore through which peptides can be co-translationally translocated into the ER lumen. Alternatively, the hydrophobic domains of transmembrane proteins can be inserted into the lipid bilayer of the ER membrane through a lateral movement out of the translocon[67] (Figure 1.4). A signal sequence peptidase complex is also associated with the translocon which cleaves the signal peptide sequence from the emerging peptide chain[64].

As the peptide chain enters the ER lumen it immediately begins to fold, with folding of individual protein domains occurring co-translationally[68]. Post-translational modifications including N-linked glycosylation occur and intra-chain disulfide bonds begin to form. Chaperone proteins are recruited to ensure correct folding of the nascent polypeptide chains, preventing aggregation and aiding isomerisation of disulfide bonds. Aside from the classical chaperones of the Hsp70 and Hsp90 families, there are also lectin chaperones (calnexin and calreticulin) that interact specifically with N-linked glycans, and chaperones that only interact with specific client proteins such as Hsp47 that is a collagen-specific chaperone.

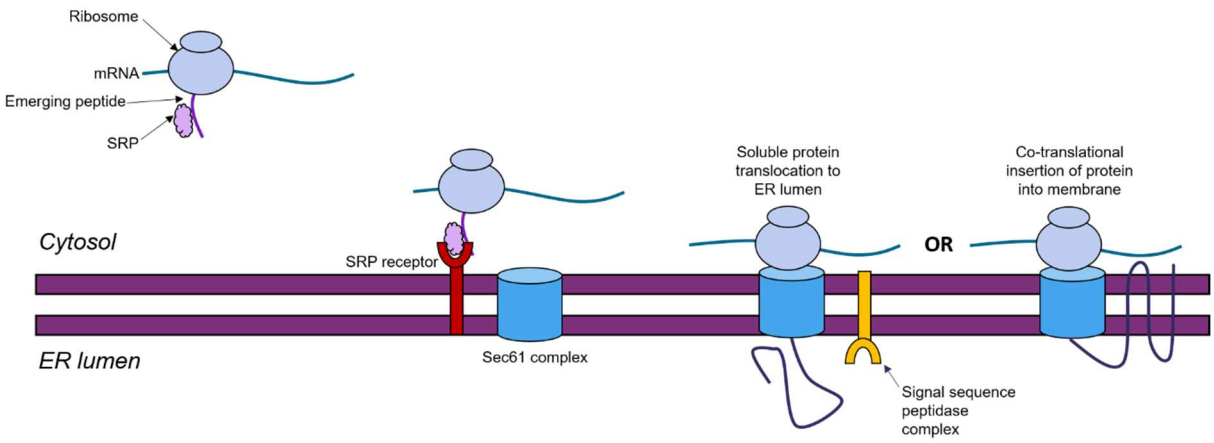


Figure 1.4: Signal peptide dependent protein targeting to the ER. The signal peptide of an emerging ER-targeted protein sequence is recognised by the signal recognition particle (SRP) causing a temporary pause in translation as the ribosome/mRNA/emerging peptide complex is docked to the Sec61 translocon by association between SRP and the SRP receptor. Once docked, translation resumes and the signal peptide is cleaved by the signal sequence peptidase complex. Proteins are co-translationally translocated to the ER lumen, or transmembrane proteins are co-translationally inserted into the ER membrane. Self-drawn adaptation from: Guerriero, C. J. & Brodsky, J. L. The delicate balance between secreted protein folding and Endoplasmic Reticulum-associated degradation in human physiology. Physiol Rev **92,** 537–576 (2012).

1.2.2.1 Classical chaperones

Hsp70 chaperones have three main domains: an N terminal ATPase domain, a substrate binding domain and a C terminal "lid". The ATPase domain has a low rate of ATP hydrolysis, and substrate affinity in the ATP-bound state is low, resulting in a high on/off rate for peptide binding and release from the chaperone. ATP activity is stimulated by co-chaperone Hsp40, leading to the hydrolysis of ATP to ADP, and a conformational change that increases Hsp70 affinity for the substrate and reduces the on/off rate. The chaperone binds to the client protein. Release of ADP by the action of a nucleotide exchange factor leads to the replacement of ADP with ATP and subsequent release of the bound peptide[69]. Hsp70 is thought to have three main roles in the ER: "holding" of unfolded proteins to prevent aggregation before necessary completion of its translation; "folding" of unfolded proteins, and re-solubilisation of misfolded protein aggregates[70]. Additional functions are found in the classical Hsp70 ER resident chaperone BiP. A BiP population resides close to the ER membrane and is thought to gate the translocon, the binding and release of the nascent polypeptide chain acting as a ratchet mechanism to prevent reverse movement of the polypeptide into the cytoplasm[71]. BiP also acts as a sensor of ER stress and is involved in the targeting of terminally misfolded proteins for ER associated degradation (ERAD) (see section 1.2.3).

Hsp90 chaperones, like Hsp70s, bind and release client proteins with concomitant ATP hydrolysis. However, unlike Hsp70s, Hsp90 chaperones function as homodimers and appear to be more selective in their protein clients, interacting with proteins that are closer to their native state, specifically those involved in signalling pathways such as the oestrogen receptor[69]. Due to this involvement in the maturation of protein clients from important cellular signalling pathways, Hsp90 chaperones have become popular targets in the development of cancer therapeutics[72]. The ER-resident Hsp90 chaperone is GRP94 and has clients including integrin proteins, insulin-like growth factor and toll-like receptors[73].

The precise molecular mechanism by which Hsp70 and Hsp90 chaperones fold their client proteins remains unclear. One suggestion is that repetitive binding and release of client proteins may actively unfold misfolded segments of the polypeptide chain to give them a second chance at folding. Alternatively, by binding client proteins it reduces the concentration of free, unfolded copies of the protein and therefore reduces aggregation, allowing the unbound copies to fold into the native state[70]. A recent study by Luengo et al., suggests that the two classes of chaperone may cooperate to stall (Hsp70) and restart (Hsp90) protein folding, not actively folding the proteins but situating them such that they can spontaneously reach their native state[74].

1.2.2.2 Lectin chaperones

Calnexin and calreticulin interact specifically with mono-glucosylated glycans. Oligosaccharides for N-linked glycosylation are built on dolichol carrier lipids on the ER membrane, initially cytoplasmic facing before "flipping" over into the ER lumen. Oligosaccharide transferase (OST) transfers the oligosaccharide to an asparagine residue onto a peptide as it exits the translocon in to the ER. Following transfer, the oligosaccharide is trimmed by α-glucosidases I and II to yield a mono-glucosylated glycan that can be recognised by the lectin chaperones (reviewed by AM Benham, 2012[64]).

Calnexin and calreticulin interact with the folding protein through a cycle of binding and release mediated by the opposing actions of glucosidase II and glucosyl transferase. Mono-glucosylated glycans are de-glucosylated by glucosidase causing the release of calnexin and calreticulin[64]. Glucosyl transferase recognises exposed hydrophobic patches of incorrectly folded proteins and re-attaches a glucose residue allowing the re-binding of the lectin chaperones. These chaperone proteins prevent the aggregation of proteins displaying exposed hydrophobic patches, and recruit protein disulfide isomerase ERp57 to catalyse the formation of disulfide bonds and promote achievement of the native state[75,76] (disulfide bond formation is discussed in more detail in section 1.2.2.4).

1.2.2.3 Client specific chaperones

There are some chaperone proteins that only act on a specific protein client. An example of particular relevance to this thesis is Hsp47 (otherwise known as SerpinH1), which binds specifically to the triple helix of procollagen and has chaperone activity[77,78]. The physiological function of Hsp47, in binding to correctly folded triple-helical structures, is not clear. Triple-helical assemblies of three procollagen peptides are stabilised by the hydroxylation of proline residues, catalysed by P4H. Distribution of stabilising residues does not occur uniformly along the length of the triple helix and the structure is therefore susceptible to local areas of relaxation[77,79]. It has been suggested by Tasab et al., that Hsp47 may function in the stabilisation of these areas that are prone to relaxation, prior to lateral association with other triple-helical procollagen units and exit from the ER[77].

Further studies by Duran et al., implicated mutations in the genes for Hsp47 and FKBP65 (another procollagen chaperone protein) in osteogenesis imperfecta, a disease characterised by low bone mineral density and brittle bones[80]. Their investigations into the effects of mutations on mouse development and collagen production in fibroblasts implicated Hsp47 not only in the stabilisation of procollagen but also in its assembly and trafficking from the ER to the Golgi. They also proposed a cooperation between the two chaperone proteins that suggests both proteins are concomitantly involved in the functional secretion of procollagen[80].

1.2.2.4 Disulfide bond formation in the ER

Proteins within the secretory pathway often contain at least one disulfide bond that forms within or between polypeptide chains. These are thought to increase the stability of the proteins within the ECM, or during exposure to acidic endocytic compartments during recycling of membrane proteins[81]. As well as being an important structural component in many secreted proteins, in some proteins disulfide bonds have also been found to serve a regulatory function. For example, redox buffers such as glutathione have been shown to regulate the function of integrins involved in thrombosis by reduction of disulfide bonds in the EGF-like domains of these proteins[82]. In the event of non-native disulfide bond formation, it is also possible for protein disulfide isomerase (PDI) to reduce these disulfides to permit re-configuration of the peptide and allow the folding protein to reach its native state[81,83] (Figure 1.5).

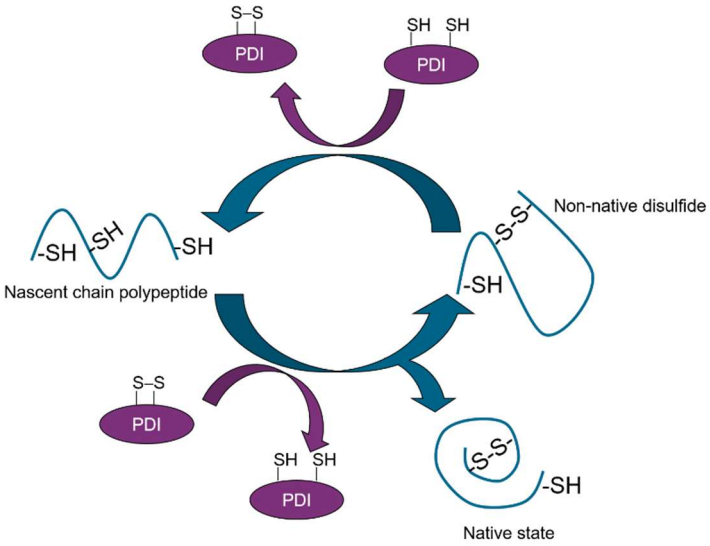


Figure 1.5: Disulfide bonds in protein folding. Protein disulfide isomerase (PDI) aids in the formation and breakage of disulfide bonds in the nascent polypeptide chain. A cycle of reduction and oxidation of the disulfide bonds continues until the protein reaches its native state or is targeted for degradation by ERAD. Self-drawn adaptation from Bulleid, N. J. Disulfide Bond Formation in the Mammalian Endoplasmic Reticulum. Cold Spring Harb Perspect Biol 4, (2012).

Disulfide bond formation is catalysed by the activity of the PDI family of proteins. Typically, PDI works in conjunction with endoplasmic reticulum oxidoreductase (Ero1α or Ero1β) to provide an electron transport chain whereby electrons are passed from client protein to PDI, to Ero1 and finally to molecular oxygen (Figure 1.6). Electrons are transferred from the client protein to PDI, resulting in the reduction of a labile disulfide bond in PDI. This disulfide is regenerated by the reduction of a shuttle disulfide bond in Ero1, which then undergoes an internal disulfide exchange to reduce another disulfide near to a bound FAD molecule. Electrons are transferred from this disulfide to FAD, reducing it to $FADH_2$. The oxidation of $FADH_2$ to FAD occurs rapidly to regenerate Ero1: $FADH_2$ transfers electrons to molecular oxygen and releases ROS in the form of hydrogen peroxide[81].

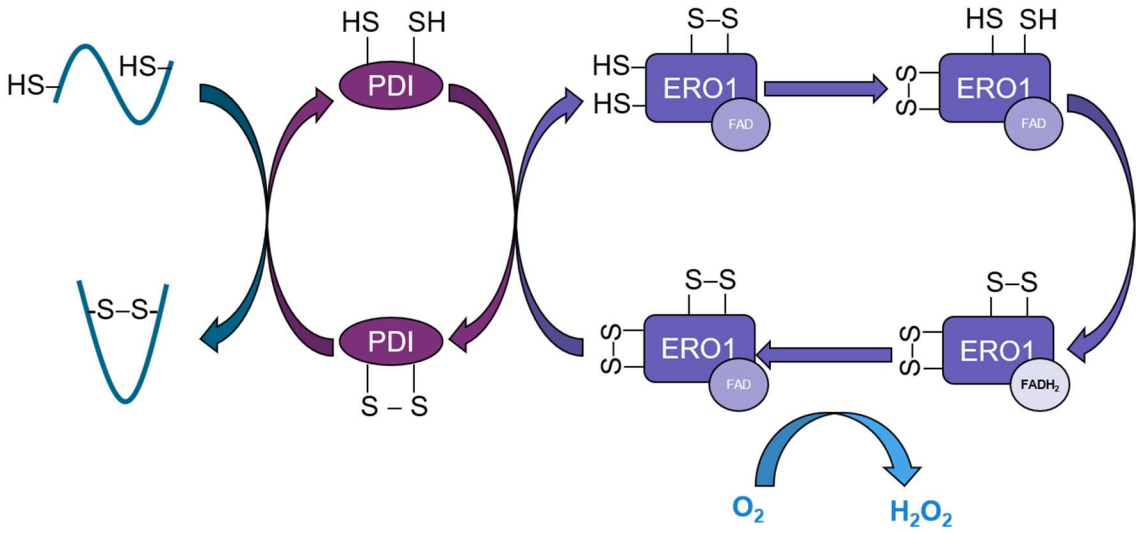


Figure 1.6 PDI and ERO in disulfide bond formation. Disulfide bond formation is catalysed by an electron transport chain from substrate polypeptide to PDI to ERO, with molecular oxygen acting as the final electron acceptor. Self-drawn adaptation from Bulleid, N. J. Disulfide Bond Formation in the Mammalian Endoplasmic Reticulum. Cold Spring Harb Perspect Biol 4, (2012).

Ero1 deletion in mice does not completely inhibit disulfide formation or result in a distinctive phenotype therefore the existence of compensatory mechanisms has been investigated[84]. In mammals, there are some alternative electron acceptors to Ero1 which can metabolise the hydrogen peroxide produced from the Ero1 cycle for example, peroxiredoxin IV (PrxIV) and glutathione peroxidases 7 and 8 (Gpx7/8)[81]. These proteins can also oxidise PDI[84]. PrxIV and Gpx7/8 use hydrogen peroxide, instead of molecular oxygen, as the terminal electron acceptor. Cysteine residues within these proteins are sulfenylated by hydrogen peroxide, before being resolved to a disulfide bond with a further cysteine in another subunit (PrxIV) or by oxidising the residues on PDI to regenerate its disulfide bond (Gpx7/8). In the case of PrxIV, this regeneration occurs from the disulfide bond between subunits re-oxidising PDI in a subsequent step[81].

There has been some debate as to whether or not disulfide bond formation requires the complete translation of protein domains or if formation is possible before this point[63]. Robinson et al., used a single domain β2M protein to demonstrate that the complete domain sequence must be exposed in the ER before formation of the disulfide bond occurred[63]. Their study indicated that PDI binds as the protein is exposed to the ER lumen, but before complete translation, and remains bound until correct intramolecular disulfide bond formation is possible, when translation has finished. This suggests that folding into native tertiary structure is important to position a cysteine residue for correct disulfide formation[63].

1.2.2.5 Peptidyl-prolyl *cis-trans* isomerases

Proline residues in the polypeptide chain are always inserted in the *trans* conformation yet frequently exist in the native state as *cis* isomers. Therefore, the activity of peptidyl-prolyl *cis-trans* isomerase (PPIase) enzymes acts to convert *trans* proline to *cis* proline to achieve the native fold[85]. This is an important and rate-limiting step in the correct folding of proteins[85]. There are three families of PPIase proteins: cyclophilins, FK binding proteins (FKBP) and parvulins. The parvulins are particularly small PPIase proteins which show tissue specific distribution and do not reside in the ER but instead function in the nucleoplasm or cytoplasm of cells[86]. Therefore, ER-resident PPIases are members of the cyclophilin or FKBP families.

Collagen, one of the key ECM proteins of the dermis, is known to require PPIase activity for efficient folding due to the large number of proline residues in the polypeptide sequence. Addition of PPIase to denatured collagen type III increased the rate of re-folding 3-fold compared to un-catalysed re-folding[87]. Additionally, treatment with cyclosporin A to inhibit cyclophilin activity greatly reduced the folding of type I procollagen in suspended chick embryo fibroblasts and furthermore increased intracellular degradation of type I and type III collagens in cultured human fibroblasts[88]. Intracellular degradation can occur as a result of proteins that are deemed unable to reach their native state as will be discussed in section 1.2.3.

Recently, the activity of FKBP65 has also been associated with the regulation of lysyl hydroxylase 2 (LH2) activity that is important for the hydroxylysine-aldehyde collagen crosslinks (HLCC) between the telopeptides of collagen molecules. LH2 is involved in the conversion of lysine to hydroxylysine in collagen telopeptides. Depletion of FKBP65 led to a decrease in the ratio of HLCCs to lysine-aldehyde collagen crosslinks (LCC), suggesting a loss of LH2 activity[89]. As LCCs are less stable than HLCCs the implication is that this results in an overall reduction in collagen fibril stability.

Once a protein has reached its native state, including the correct positioning of disulfide bonds, glycosylation and isomeration of proline residues, it can exit the ER to the Golgi by transport in COPII coated vesicles[64]. If a protein is unable to reach its native state it may be retained in the ER and targeted for degradation through the ERAD pathway. Terminal misfolding of proteins can occur due to errors in the protein sequence, insufficient amounts of an essential binding partner, overwhelming of the available chaperone machinery or perturbation of ER homeostasis[90]. The following section will discuss the quality control mechanisms by which the cell detects, retains and degrades terminally misfolded proteins.

1.2.3 Protein quality control

Due to the complex nature of protein folding and assembly in a crowded environment such as the ER, a certain level of misfolding is inevitable. In fact, it's known that a significant proportion of translated mRNAs may never reach their native state[91]. To prevent the secretion of misfolded, non-functional proteins a series of quality control mechanisms must exist to protect the reliability of the cells activities[92]. Overstretching the quality control mechanisms results in the accumulation of misfolded protein aggregates that are implicated in diseases including, for example, Parkinson's disease, Alzheimer's disease, and Huntington's disease.

A first level of quality control exists to retain incorrectly folded proteins in the ER and prevent their exit into the secretory pathway. The detection of such proteins relies solely on structural characteristics rather than a specific sequence or protein functionality. Thus, misfolded proteins that are still functional, but are perhaps likely to form aggregates, can still be retained in the ER. Such characteristics include the exposure of hydrophobic residues and unpaired cysteines[92]. These features attract the continuous binding of molecular chaperones that prevents the proteins release to the next stage of the secretory pathway. This mechanism is also affected by the stability of folding proteins. Following completion of the folding pathway, a protein in its native state may still be susceptible to local unfolding of domains before it exits the ER with, for example, fluctuations in temperature. The lower a proteins thermal stability, the more frequently this occurs and therefore the more inefficient is its secretion[92].

Glycosylated proteins are subject to a specific quality control mechanism with relation to their N-linked glycans. Proteins may not exit the ER while their glycans are still glucosylated. Removal of glucose by the glucosidase II enzyme releases the protein from the calnexin/calreticulin cycle described in section 1.2.2.2; only when this is not counteracted by the action of glucosyl transferase is the protein able to continue along the secretory pathway[92]. Glycosylated proteins that are unable to reach their native state are targeted for degradation by the trimming of mannose residues on the remaining glycan by ER resident α1,2-mannosidase I and ER degradation-enhancing 1,2-mannosidase-like protein (EDEM) enzymes[93,94]. The precise mechanism of action of each of the three known EDEM enzymes is still under review but it is thought that, in conjugation with ERdj5 (a PDI family member) and BiP (an Hsp70 chaperone), they work to prepare proteins for retrotranslocation to the cytosol and target them to ubiquitination machinery[95].

An additional level of quality control exists in the regulation of ER export of particular proteins or protein families[92]. Proteins involved in this regulation have been classified by Hermann et al., as "outfitters", "escorts" or "guides"[96]. "Outfitters" are those proteins that are involved in maintaining secretory competence prior to transport; "Escorts" are those that travel with the secretory proteins

to maintain their integrity during transport and "guides" are those that act as transport receptors on behalf of client proteins to promote their passage from the ER to the Golgi[96].

1.2.4 Endoplasmic Reticulum Associated Degradation (ERAD)

The ERAD pathway recognises and degrades terminally misfolded proteins that are unable to leave the ER through the secretory pathway. In rare cases ERAD may also accept native and functional proteins as substrates as part of a control mechanism such as feedback inhibition. For example, 3-hydroxy-3-methylglutaryl acetyl-coenzyme-A reductase (HMGR), which is an enzyme involved in sterol biosynthesis, is known to be targeted for degradation by ERAD through sterol-induced binding of the enzyme to ER membrane E3 ubiquitin ligases[97]. Ligation of ubiquitin by these enzymes to HMGR identifies it for proteasomal degradation. The ERAD pathway can also be taken advantage of by viruses for example the human immunodeficiency virus type 1 targets CD4 and major histocompatibility complex (MHC) class I proteins for degradation by ERAD to evade detection by the hosts immune system[90,98].

Ubiquitin is a small protein that plays a central role in ERAD. It is attached as a post-translational modification to proteins and targets them to the proteasomal pathway. Ubiquitin modifications are recognised by the proteasome and the conjugated protein is degraded. The role of the ubiquitin-proteasome in ERAD was first proposed in a study in yeast where a loss of function mutation in the ubiquitin conjugating enzyme Ubc6 prevented the degradation of mutant translocation apparatus[99]. Ubiquitination's role in ERAD was further established through the study of a cell culture cystic fibrosis model where degradation of CFTR was significantly reduced by inhibition of the proteasome, and a lot of the accumulated proteins were found to be ubiquitin-conjugated[90,100].

ERAD has several branching pathways that are specific to different classes of misfolded proteins[101], however they all share a common series of events: recognition, retrotranslocation, ubiquitination and finally degradation. A membrane embedded protein complex, which contains an E3 ubiquitin ligase, is involved in the coordination of these events. Research into these complexes in yeast suggests different complexes demonstrate substrate specificity dependant on the location of the misfolded domain in the protein. Three classes of complex have been established: ERAD-C, where misfolded domains exist on the cytoplasmic side of the membrane; ERAD-M, where the misfolded domains exist in the transmembrane region of the protein; and ERAD-L, where the misfolded domain is luminal[90]. However, whether similar levels of substrate specificity exist in the wider array of complexes found in mammalian cells is not yet clear[90].

Recognition of misfolded proteins has already been discussed with respect to protein quality control mechanisms in the cell (section 1.2.3). Once a misfolded protein has been recognised it

must be retrotranslocated back across the ER membrane into the cytoplasm. For soluble proteins which have been fully translocated into the ER lumen this requires a complete reversal of this process, and ubiquitination can only occur when at least a substantial proportion of the protein has been successfully retrotranslocated[90]. Conversely, with membrane associated proteins only the ER luminal and membrane domains need to be retrotranslocated and therefore ubiquitination can often be coupled to retrotranslocation. Once ubiquitinated, cytoplasmic chaperone proteins and shuttle factors maintain the solubility of these proteins and transfer them to the proteasome for degradation.

The identity of the retrotranslocation machinery is the subject of great debate in the field: Sec61, Derlin-1 and Hrd1 have all been proposed as potential candidates (reviewed by Wu and Rapoport[102]). Sec61, the translocon, had been postulated to also function in reverse however ERAD substrates were subsequently shown to interact only weakly with Sec61. In addition, the crystal structure of Sec61 revealed a plug domain across the channel that could be displaced by a growing polypeptide from the cytoplasmic side, but it is unclear whether it would be possible to displace this during retrotranslocation. Derlin-1 is a multi-spanning membrane protein that was proposed to be involved due to its interaction with p97, a cytosolic ATPase known to be critical for successful retrotranslocation of proteins. However, Hrd1 has recently been identified as another likely candidate for the retrotranslocon. Proteolysosomes containing Hrd1 were used to model protein retrotranslocation on a small scale and the autoubiquitination of Hrd1 was demonstrated to permit a misfolded luminal protein domain to cross the membrane[103]. It was therefore proposed that Hrd1 forms a ubiquitin-gated protein channel in the ER membrane through which proteins may be retrotranslocated[103].

Aggregates of misfolded proteins are inaccessible to the ERAD machinery and must therefore be cleared from the ER by an alternative mechanism. Clearance of these proteins is proposed to occur by a sequence of events similar to autophagy termed ER-phagy. Terminally misfolded protein aggregates (TMPA) are segregated to a specific area of the ER, which buds off to form a vesicle of ER membrane that contains the TMPA cargo. This vesicle is transported to the lysosome and fuses, exposing the cargo to the destructive lysosomal enzymes and completing TMPA degradation[104]. This mechanism is also thought to play a role in ER homeostasis. Under conditions of ER stress (see section 1.2.5) the ER expands, and it is thought that a similar mechanism is employed to return the ER to its pre-stress size, where vesicles of excess ER bud off and are transported to the lysosome where fusion and degradation can occur[104].

1.2.5 The Unfolded Protein Response and ER stress

ER stress can be broadly defined as conditions under which the normal physiological functions of the ER are perturbed[105]. Disruption of any of the main functions of the ER leads to ER stress. This may include inhibition of disulfide bond formation, N-linked glycosylation or the ability of chaperone proteins to aid protein folding at a more basic level. ER stress may also be induced by perturbations in lipid metabolism pathways, although the precise mechanism by which this occurs has been less well characterised[105].

The conditions of the ER are monitored by a specific set of signalling molecules that detect high levels of unfolded proteins that accumulate in the ER lumen under conditions of ER stress. Activation of these signalling components triggers a transcriptional response within the cell to downregulate general protein translation, while increasing the size and folding capacity of the ER. This response is termed the unfolded protein response (UPR). Prolonged activation of this pathway occurs when this response is insufficient to re-establish ER homeostasis and results in cell death[106].

The UPR has three main branches, named after their signalling component: inositol requiring enzyme 1 (IRE1) (Figure 1.7), double-stranded RNA-activated protein kinase (PKR)–like ER kinase (PERK) (Figure 1.8) and activating transcription factor 6 (ATF6) (Figure 1.9)[106]. All three pathways lead to the upregulation of specific UPR genes that act to increase the protein folding machinery, size of the ER, and ERAD associated proteins to clear unfolded proteins from the ER and return it to a state of homeostasis. IRE1 is the most highly conserved among eukaryotes, with PERK and ATF6 only present in the higher eukaryote metazoans[106]. There is some suggestion that, due to the number of different genes encoding ATF6 family members and IRE1 paralogs, there is a certain degree of cell and tissue type specificity in the UPR, which has not yet been fully characterised[106].

IRE1 is a type I transmembrane membrane protein that has both an ER luminal domain, and a cytosolic domain that has kinase and RNase activity[107]. Under conditions of ER stress, IRE1 monomers oligomerise via their luminal domains leading to activation and autophosphorylation of the cytosolic domains. Conformational changes that occur during this activation lead to the stimulation of RNase activity that unconventionally splices XBP1 transcripts to remove an inhibitory hairpin loop intron. Ligation of the two exons by tRNA ligase RtcB results in an mRNA transcript that produces the active XBP1 transcription factor, which moves to the nucleus and promotes transcription of UPR genes (Figure 1.7). XBP1 targets typically include ER chaperons and ERAD components such as EDEM1, but it has also been shown to act on genes involved in lipid metabolism and additional tissue-specific targets such as pro-inflammatory cytokines in macrophages[108].

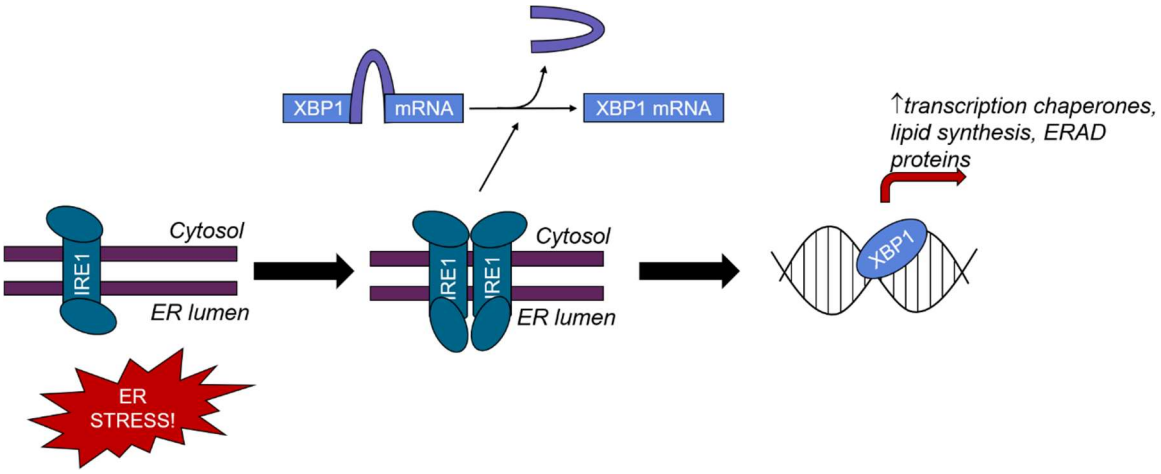


Figure 1.7: The IRE1 pathway in the unfolded protein response. Accumulation of unfolded proteins triggered by conditions of ER stress lead to the oligomerisation of IRE1, and activation of its cytosolic RNase domain. This domain splices and removes an inhibitory hairpin loop from XBP1 mRNA transcripts resulting in translation of functional XBP1 transcription factor that moves to the nucleus and promotes transcription of genes involved in chaperone production, lipid synthesis and ERAF proteins. Self-drawn adaptation from Walter, P. & Ron, D. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. Science 334, 1081–1086 (2011).

Recently, IRE1 has also been implicated in the degradation of mRNA associated with the ER (i.e. mRNA being read by ribosomes docked to the translocon for co-translational peptide translocation) called regulated IRE1-dependent decay (RIDD). A study by Tam et al., described a mechanism whereby different oligomeric states of IRE1 determined its activity[109]. Oligomeric IRE1 maintained splicing activity against XBP1, whereas IRE1 that existed only in a dimeric form acted on ER-associated mRNA as part of RIDD[109]. RIDD is thought to reduce the translational load of the ER, reducing the number of newly synthesised polypeptides that must be folded[106]. However, it has also been associated with promotion of cell death suggesting a more complex fine-tuning of IRE1 signalling that is yet to be fully understood[109].

Under normal conditions, oligomerisation of IRE1 is commonly thought to be prevented by the association of its ER luminal domain with Hsp70 chaperone BiP. During ER stress, BiP is titrated away from these signalling components by an increased load of unfolded proteins with which it interacts. This releases IRE1, allowing its oligomerisation and activation of the respective signalling pathways of the UPR (reviewed by Schröder and Kaufman[105]). The PERK pathway also involves oligomerisation of the PERK signalling component (Figure 1.8), which is inhibited by BiP association in a similar manner to IRE1.

In recent years, evidence to suggest that dissociation of BiP from IRE1 may in fact not be the trigger leading to IRE1 oligomerisation has accumulated. Crystal structures of the yeast IRE1 protein revealed a deep groove similar to the peptide-binding groove of MHC proteins, and it was therefore proposed that unfolded proteins may bind directly to IRE1, triggering its oligomerisation[107]. Although similar, the human IRE1 crystal structure did not demonstrate such a deep putative ligand-binding groove and therefore there is still some doubt as to the precise mechanism of IRE1 activation. Regardless, the overarching theory remains that an accumulation of unfolded proteins triggers the oligomerisation, and activation, of IRE1 leading to the UPR.

PERK is also a type I transmembrane protein and is structurally and functionally similar to IRE1. It has both an ER luminal stress sensing domain, and a cytosolic kinase domain that autophosphorylates upon ER stress-induced dimerisation. PERK does not, however, possess RNase activity but instead functions by phosphorylating eukaryotic Initiation Factor (eIF) 2α leading to a general reduction in protein synthesis, reducing the ER protein folding load[107]. Phosphorylated eIF2α binds nucleotide exchange factor eIF2B, stopping its activity and preventing the formation of an active eIF2 transcription complex. Although this reduces transcription of most mRNAs, the transcription of a select number of mRNAs with a specific structure is stimulated[110]. One such mRNA is that of activating transcription factor 4 (ATF4) which accumulates following PERK activation and functions to promote the transcription of UPR response genes including GADD34 and CHOP[107] (Figure 1.8).

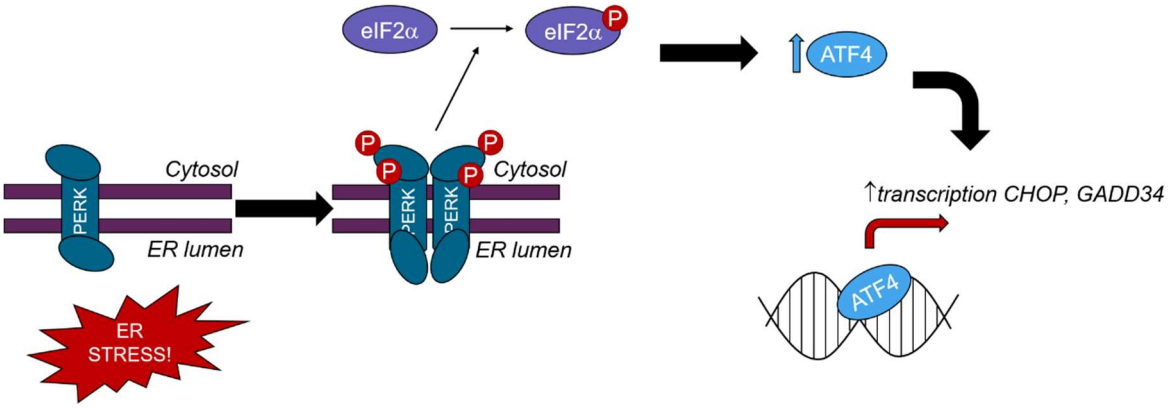


Figure 1.8: The PERK pathway in the unfolded protein response. ER stress leads to accumulation of unfolded proteins in the ER lumen. This causes dimerisation and *trans*-autophosphorylation of PERK cytosolic domain. Activated PERK phosphorylated eIF2α which reduces general protein synthesis and leads to an increase in translation of ATF4 transcription factor. Increased ATF4 promotes transcription of unfolded protein response genes, particularly CHOP and GADD34. Self-drawn adaptation from Walter, P. & Ron, D. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. Science 334, 1081–1086 (2011).

GADD34 transcription represents a negative feedback loop that acts as a pro-survival mechanism initiated by PERK. This gene encodes a regulatory subunit of a phosphatase enzyme that functions to dephosphorylate eIF2α and therefore relieve the generalised decrease in protein transcription within the cell[107]. In contrast to this, the transcription of CHOP represents a pro-apoptotic side to the PERK signalling pathway as this gene encodes a transcription factor that in turn functions to promote the transcription of pro-apoptotic factors. The opposing actions of GADD34 and CHOP demonstrates a balance point in ER stress signalling whereby at moderate levels the signalling pathways function to protect the cell and promote survival, but extended exposure to these conditions can result in the activation of apoptosis and cell death[106,107].

PERK also activates the Nrf2 transcription factor. Phosphorylation of Nrf2 by PERK releases it from the constraint of Keap1, allowing Nrf2 to travel to the nucleus where it mediates the transcription of genes regulated by the antioxidant response element (ARE). These genes include protective detoxifying enzymes such as heme oxygenase-1 (HO-1), which functions to degrade heme into bilirubin that can scavenge and detoxify peroxy radicals, and glutathione S transferase (GST), which functions to conjugate harmful electrophilic substrates to glutathione rendering them less reactive and more soluble[111,112]. Additionally, Nrf2 can also upregulate transcription of genes for chaperones proteins involved in proteasomal degradation demonstrating a role for the transcription factor in the UPR[112].

The final branch of the UPR is the ATF6 signalling pathway. ATF6 is a type II transmembrane protein, distinct from PERK and IRE1, whose cytosolic domain functions as a bZIP transcription factor. When an accumulation of unfolded proteins is detected under conditions of ER stress, ATF6 is delivered to the Golgi where the luminal domain and transmembrane anchor are sequentially cleaved by site-specific proteases. This releases the cytosolic transcription factor domain which moves to the nucleus and promotes transcription of UPR genes including chaperone proteins and components of the ERAD pathway[106,107] (Figure 1.9).

If successful, the activation of these pathways allows the cell to regain ER homeostasis and survive. However, under conditions of extreme ER stress where the UPR is not sufficient to regain homeostasis programmed cell death and apoptotic pathways are activated to remove inalterably damaged cells from the system. This process involves the activation of pro-apoptotic transcription factors such as CHOP, as described in the PERK pathway, and calcium ion (Ca^{2+}) release from the ER by activation of pro-apoptotic proteins Bak and Bax. Activated Bak and Bax proteins form a pore in the ER membrane that permits Ca^{2+} release into the cytosol. This activates calpain that converts pro-caspase 12 to caspase 12 and triggers a caspase cascade that results in apoptosis[105].

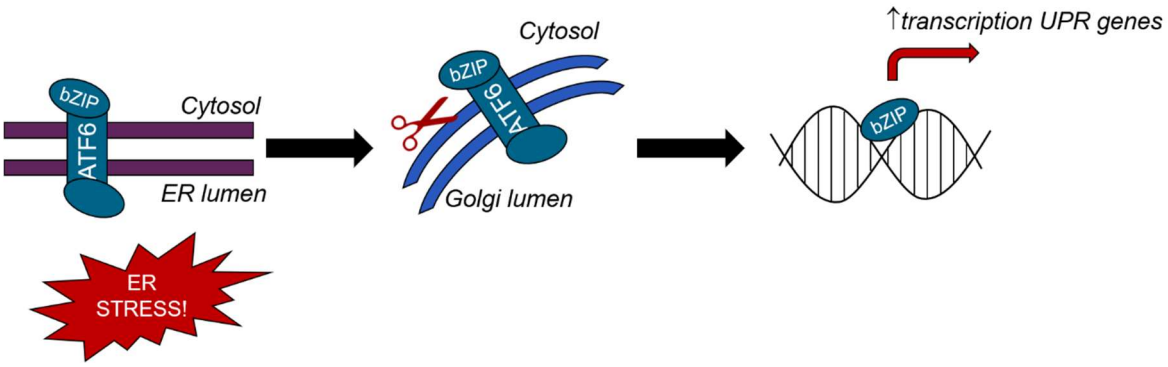


Figure 1.9: The ATF6 pathway in the unfolded protein response. Under conditions of ER stress, unfolded proteins accumulate in the ER lumen. This leads to the transport of ATF6 to the Golgi where the N-terminal domain is cleaved, releasing a bZIP transcription factor to the cytosol. The transcription factor moves to the nucleus where it promotes transcription of genes involved in the unfolded protein response. Self-drawn adaptation from Walter, P. & Ron, D. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. Science 334, 1081–1086 (2011).

1.3 Physiological impact of ER stress in dermal fibroblasts

The role of ER stress in human dermal fibroblasts is of interest due to their secretory function that suggests a high protein capacity of the ER, and the importance of protein quality control in healthy and diseased skin. For example, aged skin often suffers from fragmented collagen fibres that result in wrinkle formation and the loss of skin elasticity. Investigations into the cellular responses to ER stress may provide insight into potential therapeutic options for the treatment of damaged skin. The global anti-ageing market is estimated to be worth US$122.3 billion and is set to increase[113]; nearly a quarter of P&G's $82.6bn net sales in 2011 came from beauty products, including many anti-ageing cosmetics, demonstrating the potential impact of novel insights may have on the skincare market alone.

The effects of ER stress induced by tunicamycin and glucose starvation, which both impair N-linked glycosylation leading to an accumulation of misfolded secretory proteins, were studied by Vonk et al., to determine the effects on collagen biosynthesis both in chondrocytes (responsible for secretion of bone and cartilage ECM) and dermal fibroblasts. Quantitative PCR and immunoblotting techniques allowed the comparison of gene and protein expression of collagen, and collagen modifying enzymes and chaperones, between treated and untreated cells. This revealed a decrease in collagen production that was not associated with a similar change in expression of the collagen chaperones, suggesting the decrease in collagen production was not linked to the expression of proteins involved in its biosynthesis[114].

ER stress can be induced by several treatments, particularly those that disturb the cellular redox balance. Oxidative stress occurs when accumulation of reactive oxygen species (ROS) overwhelm the cellular anti-oxidant machinery. Disulfide bond formation in the ER is supported by the oxidising environment of the ER lumen[115]. An excess of oxidising or reductants in the ER therefore disrupts disulfide bond formation and protein folding. This demonstrates a close link between oxidative and ER stresses, whereby the redox state of the ER can be perturbed by treatment with redox agents that in turn lead ER stress and triggering of the UPR[116].

Oxidative stress also results in the oxidation of proteins, which particularly effects tyrosine residues to form tyrosyl radicals. A method to detect oxidised proteins was developed by van der Vlies et al., that harnessed the coupling ability of a fluorescein-labelled tyrosine analogue, tyramine, to proteins containing tyrosyl radicals[117]. Acetylation of the tyramine probe rendered it membrane permeable and allowed the specific labelling of oxidised proteins in response to oxidative stress[117]. This group used the acetylated tyramine probe (termed acetylTyrFluo), coupled with [^{35}S]Methionine labelling to identify proteins specifically oxidised in response to oxidative stress induced by hydrogen peroxide (H_2O_2)[118]. Protein samples from each labelling technique were pooled and separated by 2D polyacrylamide gel electrophoresis (2D-PAGE). The proteins could then

be visualised by autoradiography ([^{35}S]Methionine labelled) and western blot with an anti-fluorescein antibody (acetylTyrFluo labelled). Comparison of the results to a map of the separation of previously identified [^{35}S]Methionine labelled proteins, allowed the identification of oxidised proteins. This revealed a subset of ER resident proteins, including PDI, BiP and Calnexin, that were specific targets of oxidative stress[118] thus providing further evidence of the relationship between oxidative and ER stress.

The free radical theory of ageing dictates that ageing occurs because of aerobic metabolism that generates ROS in excess of the cellular antioxidant machinery[119]. This therefore provides a mechanism by which oxidative stress, and by association ER stress, is linked to cellular ageing. A study by Fisher et al., determined that fragmentation of collagen fibrils seen in aged skin is associated with an increased expression of matrix metalloproteinase-1 (MMP-1) that, when not inhibited, functions as a collagenase enzyme in the ECM[120]. By culturing young fibroblasts on fragmented collagen lattices, Fisher et al., described a positive feedback mechanism whereby cellular oxidation state and MMP-1 expression is elevated by fragmented collagen, which drives further collagen fragmentation of the ECM[120]. This suggests an association between oxidative stress and development of an aged skin phenotype.

A well-known cause of oxidative stress in the skin is UV-radiation from the sun. As well as causing UV-induced DNA damage, it is also known to induce H_2O_2 ROS that triggers oxidative stress within exposed cells[121]. UV radiation emitted from the sun is divided into three categories based on wavelength: UV-A (315–400 nm), UV-B (280–315 nm) and UV-C (200–280 nm). Most of the UV-C radiation is absorbed by the earth's ozone layer, leaving UV-B and UV-A waves making up the majority of UV radiation that reaches the skin. Whilst the shorter wavelength radiation UV-B is thought only to affect the cells of the epidermal skin layers, UV-A is known to reach and damage the cells of the dermis. In fact, UV-A radiation has been demonstrated to activate both IRE1 and PERK UPR pathways in normal human dermal fibroblasts[122].

Further evidence of the sensitisation of cells to oxidative stress by ECM proteins comes from Wondrak et al., who demonstrated that treatment of fibroblasts with collagen that had been pre-irradiated with UV triggered an oxidative stress response that was relieved by antioxidants[123]. Whilst *in vivo* there would never be a situation where only the ECM, and not the inhabiting cells, was exposed to UV this does suggest that ECM proteins may further promote oxidative stress in UV-exposed cells, beyond the response that might be seen in cells isolated from the ECM.

There is some evidence that provision of antioxidants can relieve the detrimental effects of oxidative stress on cells. Hydrogen peroxide (H_2O_2) treatment can be used to induce oxidative stress *in vitro* with lipid peroxidation, another side-effect of oxidative stress, often used as a tangible biomarker of its effects. Calabrese et al., demonstrated that treatment of cells with 25 µmol/L

ferulic acid ethyl ester, a nutritional antioxidant, decreased lipid peroxidation in response to 500 µmol/L H_2O_2 and promoted the expression of genes known to be involved in cellular responses to oxidative stress including HO-1 and Hsp70[124]. An additional study by Giampieri et al., determined that the antioxidant components of strawberry extract were active in protecting cells from H_2O_2, both in terms of cell viability and reduction of lipid peroxidation[125]. These groups proposed this as evidence for the protective benefit of nutritional antioxidants; however, the ability of ingested antioxidants to be absorbed in the gut and conveyed to the dermis was not addressed in these preliminary studies.

Many studies in the literature have used targeted or biased approaches to look at the effects of ER and oxidative stress on specific proteins that were thought to be of interest or had been previously identified in a mapping experiment. The development of "-omics" technologies has increased the ability to use untargeted and unbiased approaches to map global changes across an accessible genome, proteome, or metabolome. This provides opportunity to identify novel responses to treatments in biological pathways that may have previously been considered irrelevant.

The promise of proteomic techniques in the investigation of dermal fibroblast responses to oxidative stress was demonstrated by Wu et al., who used a redox sensitive 2D difference gel electrophoresis (2D DiGE) method to identify proteomic responses of skin fibroblast cells to UV-B radiation[126]. Although it is unclear whether UV-B radiation has significant physiological relevance to fibroblast cells, as it is thought most UV-B radiation from sunlight is absorbed by the epidermis, this still serves to demonstrate the power of the technique. Free thiol groups of cysteine residues not involved in disulfide bonds are vulnerable to attack by reactive oxygen species during oxidative stress, leading to the formation of sulfenic (-RSOH), sulfinic (-RSOOH) or sulfonic (-$R(=O)_2OH$) acid groups. Wu et al., used a fluorescent dye to label free cysteine groups in cell lysates. Comparison of the labelling before and after UV-B radiation allowed to identification of proteins that had been subject to thiol modification following oxidative stress. A lysine label was also used to determine changes in protein amounts independently of redox modifications. This technique identified hundreds of differentially labelled proteins from both expected and unexpected protein groups. For example, Hsp27, a chaperone that might have otherwise been thought to increase with oxidative stress was in fact shown to decrease, suggesting a role for this protein in the dermal photo-aged phenotype. Additionally proteins involved in regulating actin filament dynamics were upregulated in response to UV-B, a response that has been less well studied and may provide mechanistic insight in to the metastasis of cells that have undergone carcinogenesis following excessive UV radiation[126].

1.4 The value of proteomics

1.4.1 Defining proteomics and the importance of proteomic information

The "-omic" technologies have been developed to study the genes (genomics), which are transcribed into transcripts (transcriptomics) and then translated into proteins (proteomics) of individual cell types in different tissues and organisms. A further "-omic" approach exists in the study of metabolites (metabolomics). In an ideal world, to gain the maximum benefit from "-omic" studies, information gathered from all the technologies would be recorded and interpreted together, without prior bias or favour towards a single information source[127]. Unfortunately, such large-scale, cross-discipline studies would likely be associated with inhibitory costs, time and technology requirements for a single laboratory to perform. As such, development of complex mathematical and computational tools to integrate information, and sufficient sharing of data sets in accessible databases is being encouraged to optimise the potential benefits these investigations can have on current scientific knowledge[127]. For now, most laboratories either collaborate as consortia or focus their efforts in gaining as much insight as possible from just one or two of these "-omic" technologies.

The proteome is defined as all the protein products of the genome: it relies heavily on genomic data (see section 1.4.2) yet provides additional information that is of great importance in the study of cell and tissue behaviour[128]. Whilst genomic and transcriptomic data may provide insight into DNA damage and regulation of gene expression in response to environmental stimuli, the protein copy number within a cell often differs markedly from the transcript copy number. In addition, proteins may be regulated by post-translational modifications and therefore a change in protein activity cannot always be directly implied from a change in gene expression. Proteomic techniques enable investigation into changes in protein abundance in response to a stimulus, as well as facilitating the detection of post-translational modifications and protein-protein interactions that shed more light on cellular responses than can be gleaned from genomic and transcriptomic approaches alone.

1.4.2 Proteomic methodology

Mass spectrometry has become the main method by which scientists investigate the proteome. This methodology indirectly relies on gene sequence data, which can, in turn, provide information on the amino acid sequences of known gene products. In any organism, most amino acid sequences greater than 6 residues are unique. That is, the same 6 amino acid sequence is unlikely to be found in the sequence of another gene product, unless they are closely related and therefore the sequences highly conserved[129]. This means that if the sequence of at least one or two peptides from a protein can be confidently identified, it is possible to identify the protein from which they came. To do this, peptides must be generated from proteins in the sample before their sequences can be determined using mass spectrometry. As peptide sequences are usually determined by comparing the measured mass of peptides to the mass of predicted peptides based on the protein sequence, the proteolytic enzyme used must have predictable cleavage sites. This allows the *in-silico* determination of cleavage sites in protein sequences, and prediction of peptide masses, for comparison to measured data. Usually, proteins are digested into peptides with trypsin, a protease enzyme whose cleavage sites are specific and predictable, namely C-terminal to arginine and lysine residues.

Mass spectrometers are made of three main components: the ion source, mass analyser, and detector. Peptides must be in the gas phase to be analysed, therefore the ion source serves to ionise peptides. Peptides are large, polar molecules which are not easily ionised without fragmentation or decomposition[130]. Electrospray ionisation (ESI) and matrix assisted laser desorption/ionisation (MALDI) are termed “soft” ionisation techniques and they have enabled the ionisation of these molecules without damage[131,132]. In the liquid chromatography-coupled mass spectrometry, to be employed in this thesis, electrospray ionisation is used. This works by passing a sample solution through a narrow, electrically charged needle tip to release small droplets carrying a positive charge. As these droplets spray from the needle, solvent evaporation leads to droplet disintegration, increasing the surface charge to a maximum at which point the ions are emitted into the gas phase[133].

Once in the gas phase, ions can be analysed in the mass analyser. Several different types of mass analyser are currently in use including, for example, the time-of-flight (TOF), quadrupole, ion trap and Fourier transform ion cyclotron (FT-MS) analysers. For peptide analysis, tandem mass spectrometry is usually employed whereby the mass of the parent (or precursor) ion is first measured before fragmentation in a collision cell and then masses of the resulting daughter (or fragment) ions are measured in a second mass analyser. The machines used for the analysis presented in this thesis have two mass detectors: both a quadrupole, and a time of flight. In this set up, a quadrupole selects ions with a specific mass to charge ratio (m/z) which then enter a collision

cell for fragmentation before measurement on a final TOF analyser[134]. The m/z value (dimensionless) is the unit displayed on the x-axis of mass spectra and is determined by the molecular mass divided by the charge of a given ion.

Quadrupoles are made up of four parallel rods, through which ions travel. An electrical field is applied to the rods consisting of a constant voltage, and a radiofrequency voltage. This voltage causes ions to fly in a corkscrew-like flight path, which will only be stable with ions of a specific, and narrow, mass range for any given voltage setting. Ions which are not within the permitted m/z range spin out and hit the rods, whereas ions that are can fly through the rods and reach the detector or collision cell[135]. A TOF analyser works on a much simpler principle: that if ions of the same charge are accelerated by the same force, smaller ions will fly faster than bigger ions. Thus, by measuring the time taken for an ion to travel through a flight tube and hit a detector, its size can be calculated[135]. The detector is the final component of the mass spectrometer, and this measures the intensity of the signal, which is relative to the number of ions at the m/z value.

Tandem mass spectrometry produces two sets of spectra: first the MS1 data which provides information on the weight and intensity of the parent ions, and second the MS2 data, which provides information on the weight and intensity of the daughter ions. Peptides fragment in a predictable manner and it is therefore possible to use this information to determine the sequence of amino acids in the peptides identified. When peptide ions collide with the gas ions in the collision cell, the bonds along the backbone break to produce “y” and “b” ions (Figure 1.10). The “b” ions are the N-terminal side of the peptide, with a positive charge on the carbon of the carboxyl group; the “y” ions are the C-terminal side of the peptide, with a positive charge on the nitrogen of the amino group. With efficient ionisation of peptide parent ions, it is possible to see a complete series of y and b ions in the MS2 spectra. Comparing the masses identified to the known masses of amino acids, it is then possible to determine the sequence(s) of the peptide(s) identified in MS1.

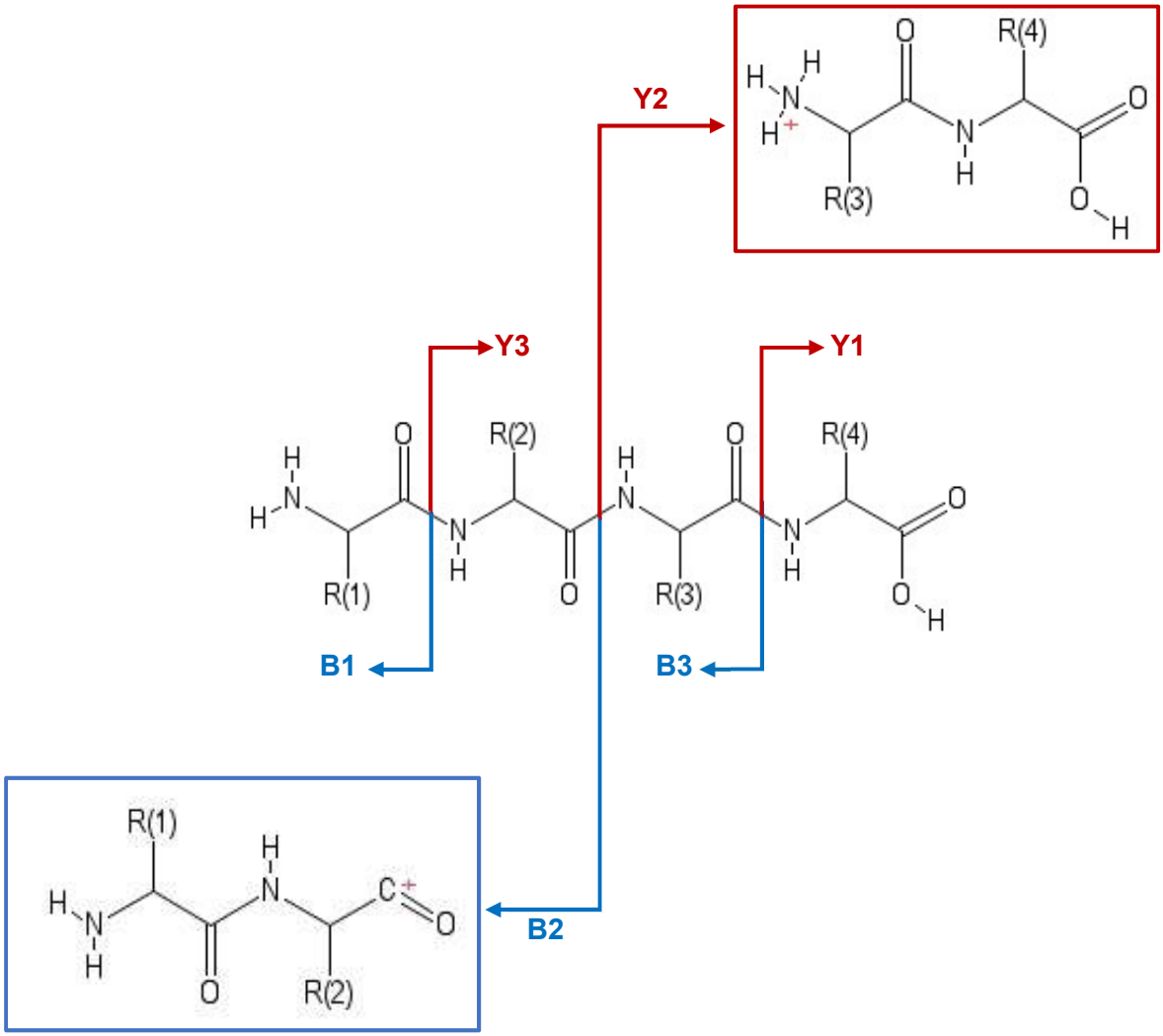


Figure 1.10: Peptide ionisation into y and b ions. When peptides fragment, the peptide bond breaks to generate b ions, containing the now positively charged carboxyl group, and y ions containing the now positively charged amide group. A well ionised peptide should produce a series of y and b ions, where all or nearly all the peptide bonds are broken to generate smaller and smaller fragment ions.

When analysing proteins in cell lysates, the resulting spectra consist of daughter ions from several parent peptides. These spectra are often too complex to identify the parent peptides manually, and thus these are analysed *in silico* by searching databases of known fragmentation patterns. To fully analyse such a complex sample of proteins, it is usually necessary to separate the mixture e.g. by strong cation exchange chromatography, to reduce complexity, and therefore maximise the number of identifications seen. If the sample becomes too complex, it increases the chance of under sampling where the same proteins aren't identified in repeats[136]. Fractionation also improves the dynamic range of analysis, allowing for a greater range in protein abundances to be identified. In an unfractionated sample, only those parent ions with the greatest intensity (and therefore highest abundance in the sample) will be selected for fractionation. By separating the mix of peptides, the intensity threshold that must be met for parent ions to be fragmented is reduced. The identifications derived from spectra from all the fractions can be combined to create a fuller picture of the samples proteome.

Fractionation can use either gel-free or gel-based techniques. Proteins can be separated using 2D-PAGE, which resolves proteins into distinct spots on the gel, based on isoelectric point and molecular weight. These spots can then be cut out of the gel and proteins extracted and digested for mass spectrometry analysis. Alternatively, proteins or peptides can be separated by liquid chromatography, which resolves them based on their chemistry. For example, strong cation exchange (SCX) is commonly used to separate peptides based on their charge. SCX material usually contains many sulfonic acid groups, which are negatively charged in solution, and interact with positively charged cations. Washing with increasingly basic solvent will gradually elute peptides by neutralising these ionic interactions[137]. SCX will be used for fractionation of relevant samples in this thesis (See Chapter 3).

1.5 Thesis aims

Although there is a clear link between ER stress and oxidative stress, and several proteomic responses have been determined in dermal fibroblasts, relatively little is known about the effects of perturbing cellular redox balance in the opposite direction, namely, how reductants affect the human dermal fibroblast proteome. The skin may be exposed to reductants from a range of sources: from sulfur dioxide in the atmosphere released from burning fossil fuels to thioglycolate used in cosmetic depilatory creams. This thesis aims to investigate the proteomic effects of reductive stress on human dermal fibroblasts by:

1. Developing tools, protocols and procedures for the identification of protein responses to environmental stimuli from dermal fibroblast cell lysates and conditioned cell growth media.
2. Using a range of biochemical and proteomic techniques to investigate the dermal fibroblast response to reductive stress.
3. Using the same methodology to profile the proteomic response of these cells to oxidative stress induced by UV-A radiation and comparing this to reductive stress.

Chapter 2: Materials and Methods

2 Materials and Methods

All chemicals used in the work pertaining to this thesis were obtained from Sigma Aldrich unless otherwise stated. Where catalogue numbers are given without a company descriptor these correspond to Sigma products.

2.1 Cell culture

Cultures of human dermal fibroblasts were maintained in a humidified incubator at 37 °C and 5% CO_2. Adherent fibroblasts were sub-cultured twice a week in T75 flasks (TPP, 90076) and passaged under sterile conditions in a CL2 flow hood (SteriGUARD, The Baker company, Sanford, Maine 04073 USA). BJ fibroblast cells (ATCC® CRL-2522™) and normal human dermal fibroblasts (nHDF) (neonatal) (gifted by Pamela Ritchie, Durham University) were grown in Eagle's Minimum Essential Medium (MEM) (Gibco, 41966-029) supplemented with 100 units/ml Penicillin, 100 µg/ml Streptomycin (Gibco 15410), glutamax (Gibco, 35050-038) and 10% Foetal calf serum (FCS) (F7524). 142BR fibroblast cells (ECACC 90011806) were grown in MEM (Gibco, 41966-029) supplemented with 100 units/ml Penicillin, 100 µg/ml Streptomycin (Gibco 15410), glutamax (Gibco, 35050-038), 1% non-essential amino acids and 15% FCS (F7524). Prior to passaging, the medium was warmed to 37°C for 15 mins in a water bath. During this time, the spent media from the flask was removed and the cells were washed twice with 10 ml sterile phosphate buffered saline (PBS; 0.2 g/L KCl, 0.2 g/L KH_2PO_4, 8 g/L NaCl, 1.15 g/L Na_2HPO_4 (anhydrous), D8537). 1 ml of trypsin (0.05% Gibco, 15400-054) was applied to the cells for 1 min, 0.5 ml was then removed, and cells were incubated at 37 °C for five minutes. A single cell suspension was created by adding an appropriate volume of media (10 ml for a T75 flask) and aspirating. Fresh media was added to a new sterile T75 flask sufficient to create the desired dilution (5 ml for a standard 1:2 passage) and then the corresponding volume of cells (5 ml for a standard 1:2 passage) added and mixed by aspirating. The sub-cultured cells were returned to the incubator to incubate for 3-4 days.

2.2 Cell lysis

Cells grown to 90% confluence were washed twice with PBS (D8537), and then placed on ice after aspiration of the PBS. To solubilise ER and cytosolic proteins, 300 µl (T25 flask) or 600 µl (T75 flask) MNT lysis buffer was added to the cells. 1x MNT buffer was comprised of 20 mM MES (M8250), 30 mM Tris-HCl (T3253), 100 mM NaCl (S7653), pH 7.4, with 1% v/v Triton X-100 (X100) and 10 µg/ml each of protease inhibitors chymostatin, leupeptin, antipain, pepstatin A. The cells were left to lyse on ice for 20 minutes prior to scraping thoroughly with a cell lifter (TPP, 99003T). The resulting lysate was then transferred to a 1.5 ml Eppendorf tube, vortexed and centrifuged at 16 000 g, 4 °C for 10 minutes to pellet the nuclei, and the post nuclear supernatant transferred to a fresh tube.

Pellets were discarded or retained if required. Lysates were flash frozen in liquid nitrogen and stored at -20 °C for further analysis.

Alternatively, cells were lysed in RIPA buffer to retain nuclear proteins. 600 µl RIPA lysis buffer (1% v/v Triton X-100 (X100), 50 mM Tris HCl (T3253), pH 8, 150 mM NaCl (S7653), 0.5% w/v Na-deoxycholate (D6750), 0.1% w/v SDS (74255)) supplemented with 10 µg/ml each of protease inhibitors antipain, chymostatin, leupeptin and pepstatin A and 1x phosSTOP from a 10x stock of 1 tablet in 1ml dH2O (against acid, alkaline, serine/threonine, tyrosine and dual-specificity phosphatases; Roche 04906845001) was used per T75 flask of PBS-washed cells as described above. The cells were left to lyse at room temperature to avoid precipitation of SDS for 1 minute prior to scraping thoroughly with a cell lifter. Lysates were transferred to a 1.5 ml Eppendorf tube and vortexed for 30 seconds before being centrifuged at 16 000 g, 4 °C for 10 minutes to clear any debris from the sample. Supernatants were transferred to a fresh tube and pellets discarded. Lysates were flash frozen in liquid nitrogen and stored at -20 °C.

2.3 Cell treatments

2.3.1 Chemical treatments

Unless otherwise stated, cells were treated in serum free MEM. Cells were grown to 70% confluency, washed once with PBS and incubated with serum free media. Chemicals were added to media to give final concentrations of 5 ng/ml platelet derived growth factor (PDGF) -bb (Gibco, PH90045), 5 mM dithiothreitol (DTT) (D9779), 2 µM Thapsigargin (T9033), 12.5 µM Etoposide (E1383), 5 mM Peroxide (H1009), 5 mM Taurine (T8691), 5 mM Potassium Thioglycolate (Acros organics, 384702500). Cells were kept at 37 °C in a humidified incubator for the duration of treatment, before washing and lysing as described above.

2.3.2 UV radiation

Cells were grown to 70 % confluency, spent media set aside and cells washed once with PBS. All fluid was removed, and control cells wrapped in foil before inversion on a transilluminator set to an intensity of 8000 µW/cm^2 for 10 minutes (to give final dose of 5 J/cm^2). Spent media was then replaced and cells returned to a humidified incubator at 37 °C to recover for 18 hours before lysis in RIPA buffer as described above (section 2.2).

2.4 Orangu™ cell proliferation assay

Cell proliferation was measured using the Orangu™ cell counting solution (Cell Guidance systems, OR01-500) according to manufacturer's instructions. Briefly, cells were seeded into a 96-well plate and allowed to settle overnight. Following treatment, cells were washed and replaced in to fresh media. On each day that readings were taken, 10 µl Orangu™ reagent was added and incubated for 2 hours before reading at 450 nm on a plate reader (BioTek, ELx800 Absorbance Reader).

2.5 Tissue sample lysis

Murine liver tissue was cut into small pieces and snap frozen in liquid nitrogen until use. After defrosting, tissue was cut into small pieces and pushed through a 40 µm cell strainer (BD biosciences) with 1X MNT lysis buffer. The resulting lysate was vortexed and then centrifuged at 16 000 g, 4 °C for 10 minutes. The supernatant was transferred to a fresh tube and the pellet stored at -20 °C. The protein concentration of the tissue lysate supernatant was estimated by Bradford assay (section 2.6.1) and diluted to ~10 µg/µl before use.

2.6 Protein concentration estimation

2.6.1 Bradford Assay

A standard curve was prepared using Bovine Serum Albumin (BSA) (A9647) at concentrations of 0, 1, 2, 5, 8, and 10 µg protein by dilution of 1 mg/ml BSA in lysis buffer corresponding to the samples being measured and each mixed with 10 µl 0.1 M HCl, 80 µl H_2O and 900 µl Bradford dye (BioRad, Hercules, CA, USA). Samples to be tested were prepared with 2 µl sample in 8 µl lysis buffer, with 10 µl 0.1 M HCl, 80 µl H2O and 900 µl Bradford dye (BioRad, Hercules, CA, USA). After addition of dye, samples and standards were incubated at room temperature for 10 minutes before the absorbance was read at 595 nm using an Eppendorf Biophotometer (Eppendorf AG, Hamburg, Germany). A standard curve of protein concentration against absorbance was plotted and used to estimate protein concentrations from the unknown samples.

2.6.2 BCA Assay

The BCA assay is thought to be more detergent compatible than the Bradford assay, and was used for protein estimation of samples being prepared for mass spectrometry analysis by the TripleTOF 6600 mass spectrometer. The BCA assay was performed using the Peirce™ BCA Protein Assay Kit (ThermoFisher Scientific, 23225) according to kit instructions in a microplate. Briefly, standards at 0, 25, 125, 250, 500, 750, 1000, 1500, and 2000 µg/ml were prepared from a 2 mg/ml BSA stock. Aliquots of 25 µl from each standard and sample being tested were added to wells of a 96-well flat-bottomed microplate in duplicates. 200 µl working reagent was added to each well and incubated at 37 °C for 30 minutes before reading in a plate reader at 562 nm. Standards were used

to plot a standard curve of protein concentration against absorbance, from which the protein concentrations of unknown samples could be measured.

2.7 SDS PAGE

SDS PAGE 10% gels (6 ml dH_2O, 4.95 ml 30% Acrylamide, 3.75 ml 1.5 M Tris pH 8.8, 150 µl 10% SDS, 150 µl 10% APS, 6 µl TEMED) were cast using the Hoefer® Dual Gel Caster for mini gels (Hoefer Inc., USA), with a stacking gel (1.42 ml dH_2O, 0.33 ml 30% Acrylamide, 0.25 ml 1 M Tris pH 6.8, 20 µl 10% SDS, 15 µl 10% APS, 2 µl TEMED) containing wells for protein loading cast on top. A maximum volume of 15 µl of sample mixed 1:1 with 2 x Laemmli sample loading buffer containing 65.8 mM Tris-HCl, pH 6.8, 26.3% (w/v) glycerol, 2.1% SDS and 0.01% bromophenol blue (Biorad, #1610737) and reduced with 50 mM DTT before loading against a stained or unstained marker, depending on future work to be carried out with the gel. The gel was run in a tank containing 1 x Tris-Glycine electrophoresis buffer containing 25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3 (diluted from 10 x stock, BioRad, #1610772) at between 10 and 50 mA. Gels were then stained with colloidal Coomassie (section 2.8) or transferred for western blotting (section 2.9).

2.8 Coomassie Staining

Gels were fixed in 1x fixing solution containing 12% trichloroacetic acid and 3.5% 5-sulfosalicyclic acid (diluted from 5 x stock, F7264) for 10 minutes with gentle shaking before being transferred to colloidal Coomassie solution containing 0.1% (w/v) Brilliant Blue G, 0.29 M phosphoric acid and 16% saturated ammonium sulfate (B2025) for 1.5 hours or overnight. After staining gels were placed in the first de-stain solution (5 ml acetic acid, 12.5 ml methanol, 32.5 ml distilled water) for 10 minutes and finally transferred to 25% methanol destain solution. The gel was incubated at room temperature with gentle rocking until such a time as the gel was analysed by scanning, with regular refreshing of the final de-stain solution until the excess blue colour was removed.

2.9 Western blotting

Proteins separated by SDS-PAGE gel were transferred to a methanol-primed PVDF membrane (Immobilon) using the BioRad Mini Trans-blot® system. Transfer was performed at 4 °C in transfer buffer (25 mM Tris, 1.5 M Glycine and 20% methanol) for 2.5 hours at 150 mA or overnight at 30 V. After transfer, membranes were blocked in 5 % milk in Tris buffered saline with Tween (TBS-T) (10 mM Tris pH 8, 150 mM NaCl, 0.1% Tween-20 (P1379)) for 1 hour at room temperature before incubation with primary antibody at 4 °C overnight. Primary antibodies were diluted according to dilutions detailed in Table 2.1, either in 1 % milk in TBS-T (PDI, pan-Akt, β-actin) or 5% BSA in TBS-T (all other antibodies listed in Table 2.1). After antibody incubation, membranes were washed for 5 minutes in TBS-T 5 times and then incubated with secondary antibody for 1 hour at room

temperature. HRP-conjugated secondary antibodies were diluted in 1% milk in TBS-T at 1:3000. Swine anti rabbit (SARPO: Dako, P0217) or goat anti mouse (GAMPO: Dako, P0447) were used depending on animal of origin (see Table 2.1). Membranes were washed a final 5 times in TBS-T as before, prior to visualisation of target proteins with ECL (GE Healthcare, RPN2232) and light sensitive film.

2.10 Immunoprecipitation

A 200 µl aliquot of cell lysate was incubated with 0.5 µl antibody against target overnight at 4 °C on a rotor. A 100 µl volume of sepharose beads coated with protein A was then added and the samples and incubated for a further 4 hours at 4 °C. The sepharose beads were then pelleted by spinning at 1500 g and the supernatant removed. The pellet was washed 2 times with lysis buffer, and then a further 4 times with 1 x MNT buffer before elution with 50 mM NH_4HCO_3, 50 mM DTT, 1% SDS by boiling for 5 minutes. The eluate was collected and analysed by Bradford assay to determine protein content. Samples were then taken forward for analysis by in-solution digestion and mass spectrometry as previously described.

2.11 Immunofluorescence

Cells for analysis by immunofluorescence were seeded onto 18 mm coverslips in 6 cm dishes before treatment. Treatments were performed once cells had reached suitable confluency on these cover slips, after which the cells were washed twice with PBS containing 1 mM $CaCl_2$ and 0.5 mM $MgCl_2$ (PBS^{++}). Cells were fixed for 15 minutes in 4% paraformaldehyde, washed again in PBS^{++} and the free aldehyde groups quenched with 50 mM NH_4Cl_2 by incubation for 10 minutes. Cells were then permeabilised using 0.1% Triton X-100 (X100), washed once for 5 minutes with PBS^{++} and a further two times, for 5 minutes each, with 0.2% BSA in PBS^{++}. Coverslips were inverted onto parafilm for incubation with 20 µl anti-PDI antibody diluted appropriately with 0.2% BSA in PBS^{++} (see Table 2.1) for 20 minutes, protected from the light. Following primary antibody incubation, the coverslips were returned to 6 cm dishes for three, 5-minute washes with 0.2% BSA in PBS^{++}. The coverslips were then inverted on parafilm again for incubation with 2 µg/ml fluorescent conjugated donkey anti-rabbit AlexaFluor 594 secondary antibody (Invitrogen, A21207). Finally, coverslips were washed twice with 0.2% BSA in PBS^{++} before incubation with 50 µl 1x DAPI (Invitrogen) for 5 minutes in the dark. Coverslips were rinsed briefly in dH_2O to remove excess DAPI and PBS^{++} before mounting onto microscope slides with Vectashield (Vector, H-1400) and sealing with clear nail varnish. Slides were imaged using an inverted wide field fluorescence microscope (Zeiss Apotome).

2.12 Antibodies

All antibodies used were commercially available except for anti-PDI and anti-Ero1α. The polyclonal anti-sera against PDI was generated by immunising rabbits with a purified rat PDI protein preparation as described by Benham et al[138]. The 2G4 antibody against Ero1α was a kind gift from Professor Roberto Sitia. The antibodies, their application (use), supplier and dilutions are detailed in Table 2.1.

Table 2.1: Antibodies. The type, usage and supplier for all antibodies used in this thesis.

Antibody	Animal/Type	Use	Dilution	Supplier/Catalogue Number
β-Actin	Mouse mAb	WB	1:10000	Abcam #8224
PDI	Rabbit pAb	IF	1:200	Benham laboratory (Benham et al., 2000[138])
phospho-Akt (Ser473)	Rabbit pAb	WB	1:1000	Cell Signalling Technology® #9271
phospho-ERK1/2 (Thr202/Tyr204)	Rabbit mAb	WB	1:2000	Cell Signalling Technology® #4370
PathScan® PDGFR Activity Assay Multiplex Western Detection Cocktail II	Rabbit pAb	WB	1:200	Cell Signalling Technology® #5304
phospho-eIF2α	Rabbit pAb	WB	1:500	Cell Signalling Technology® #9721
pan-Akt	Mouse mAb	WB	1:2000	Cell Signalling Technology® #2920
Ero1α (2G4)	Mouse mAb	IP	1:2500	Kind gift from Prof. Roberta Sitia

2.13 Concanavalin A

An aliquot of 50 µl concanavalin A-coated sepharose beads were washed and resuspended in lysis buffer to 70% solution. Approximately 50 µl of this bead solution was then added to 500 µl lysate and incubated on rotor at 4 °C for 30 minutes. Following incubation, the sample was centrifuged for 1 minute at 6000 rpm and the supernatant removed. The remaining beads were washed twice with lysis buffer before re-suspending in 2 x Laemmli sample loading buffer containing 65.8 mM Tris-HCl, pH 6.8, 26.3% (w/v) glycerol, 2.1% SDS and 0.01% bromophenol blue (Biorad, #1610737), boiled for 3 minutes and centrifuged at 13 000 rpm to pellet beads. Supernatants were loaded onto 10% SDS-PAGE gels for analysis.

2.14 2D PAGE

Protein samples for 2D PAGE were prepared in a buffer of 9M Urea, 2M Thiourea, 4% CHAPS either by cell lysis directly into the buffer or by acetone precipitation of protein sample and re-suspension. Immobiline DryStrips (7cm pH 3-11 non-linear; GE Healthcare) were set to re-swell overnight in a groove of a re-swelling tray containing 100 µg protein sample to be separated, 2.5 µl IPG buffer, 2.5 µl 50% DTT, 20 µl lysis buffer spiked with bromophenol blue, and sufficient lysis buffer to make this up to a volume of 125 µl. The strips were covered in Immobiline™ DryStrip Cover Fluid (GE Healthcare) during re-swelling and left on a level surface overnight. After re-swelling, the strips were removed, rinsed in water and blotted on damp filter paper before loading onto Multiphor™ II Electrophoresis system (GE Healthcare) and connected to the MultiTemp™ III thermostatic circulator for isoelectric focussing (IEF). Two wicks soaked in water were placed at either end of the gel strip, overlapping the gel and backing strip. Focussing was run in three phases: 200 V, 1 mA, 5 W, 10 kWh; 3500 V, 1 mA, 5 W, 2700 kWh; 3500 V, 1 mA, 5 W, 2800 kWh. Following isoelectric focussing, the IEF strips were equilibrated in buffer (50 mM TrisCl pH 8.8, 6 M Urea, 30% Glycerol, 10% SDS) with 10 mg/ml DTT for 15 minutes, and then again in equilibration buffer with 48 mg/ml iodoacetamide for a further 15 minutes. Following equilibration, the gel strips were loaded onto 10% SDS PAGE resolving gels with a molecular weight marker and sealed using 0.5% agarose spiked with bromophenol blue. The second dimension resolving gels were run at 30 mA before Coomassie staining as described above (section 2.8).

2.15 Senescence-Associated Beta Galactosidase assay

Staining for senescence-associate β-galactosidase activity was performed using the Senescence β-Galactosidase Staining Kit from Cell Signalling technology® (9860) according to manufacturer's instructions. Briefly, spent media from treated cells was removed and cells washed once with PBS before fixing in a proprietary fixative solution for 10-15 minutes at room temperature. Cells were then washed a further two times before application of pH-adjusted β-galactosidase staining

solution and incubation overnight in a dry incubator at 37 °C. Staining solution was then removed, and cells washed twice with PBS. Following staining, dishes of stained cells were either stored in glycerol at 4 °C or cover slips DAPI stained and mounted on microscope slides for analysis. Images were taken using a CMEX WiFi 5 camera (Euromex, DC.5000-WIFI) attached to an inverted light microscope or inverted wide field fluorescence microscope for DAPI stained cells (Zeiss Apotome).

2.16 RNA extraction and RT-PCR

Following two PBS washes, cell lysates for RT-PCR analysis were obtained using 250 µl TRI reagent (T9424) and scraping with a cell lifter before transfer to Eppendorf tubes. Cell lysates were incubated for 10 minutes on ice with 50 µl chloroform (Fisher Bioreagents, BP1145-1) after vortexing. Lysates were then centrifuged at 16 000 g for 15 minutes at 4 °C and the aqueous phase removed to a fresh Eppendorf tube. 125 µl isopropanol (BDH, 1525265) was added to this aqueous phase, vortexed and incubated on ice for a further 10 minutes. Samples were again centrifuged for 10 minutes at 16 000 g, 4 °C, supernatants discarded, and the pellets washed with 75% ethanol. Air-dried washed pellets were re-suspended in nuclease-free water to a final concentration of 50 ng/µl (concentrations were determined using a nanodrop spectrophotometer (ND-1000 LabTech international)).

RT-PCR was performed using the AccessQuick™ RT-PCR system from Promega (A1701) according to the manufacturer's instructions. The primers for β-actin were: CCACACCTTCTACAATGAGC (forward) and ACTCCTGCTTGCTGATCCAC (reverse) and the primers for XBP1 were: GAAACTGAAAAACAGAGTAGCAGC (forward) and GCTTCCAGCTTGGCTGATG (reverse). The protocol used in the thermal cycler was: 1 hour at 45 °C; 2 minutes at 94 °C; 30 cycles of {30 seconds at 94 °C, 1 minute at 60 °C, 1 minute at 72 °C}; 5 minutes at 72 °C and finally reduced to 4 °C where the temperature was maintained until retrieval of samples. XBP1 products were subject to digestion with 10 U/µl *PstI* (Thermo, ERO611) for 2 hours by incubation in 37 °C waterbath. The products were cleaned up using a PCR purification kit from Qiagen (QIAquick, Qiagen, 28104) before analysis. RT-PCR products were analysed on a 1% (β-actin) or 2% (XBP1) agarose gel. Agarose (Melford, MB1200) was dissolved at the appropriate w/v ratio in 0.5% TAE buffer with 5.5 µl ethidium bromide. Gels were run using a Bio-Rad Sub-Cell® GT Horizontal Electrophoresis System (#1704401) at 100 mV for 40 minutes before visualising under UV light.

2.17 Liquid chromatography – tandem mass spectrometry (LC-MS/MS) Sample preparation

2.17.1 In-gel digests

Gel spots for protein identification were taken using a pipette tip and transferred to an Eppendorf tube containing dH_2O until de-stained. The water was then removed, and the gel spots incubated 3 times for 10 minutes in 100 mM ammonium bicarbonate and 40% v/v acetonitrile. Following aspiration of the final ammonium bicarbonate/acetonitrile wash, the gel spots were incubated for 10 minutes in HPLC grade acetonitrile and then 10 minutes in HPLC grade H_2O; these two incubations were repeated sequentially twice before a final 10-minute incubation with HPLC grade acetonitrile. The spots were then vacuum dried for 10 minutes at 60 °C before adding 50 µl 40 mM ammonium bicarbonate, 9% v/v acetonitrile with 20 ng/µl sequencing grade trypsin (Promega, V511A) for 30 minutes. A further 200 µl 40 mM ammonium bicarbonate, 9%v/v acetonitrile was then added, and spots incubated for 16 hours at 37 °C before acidification with 25 µl 1% trifluoroacetic acid (TFA). Samples were then taken for mass spectrometry analysis as described below (section 2.18).

2.17.2 In-solution digest

Samples for in-solution digestion were incubated with 0.01 volumes of 0.5 M DTT for 20 minutes at 56 °C, cooled to room temperature and then incubated with 0.025 volumes 0.5 M iodoacetamide for 20 minutes at room temperature in the dark. Modified sequencing grade trypsin (Promega, V511A), re-suspended in 10% acetonitrile, 2 mM HCl, was added at a ratio of 20-40:1 w/w protein:enzyme and incubated at 37 °C overnight.

2.17.3 Filter Aided Sample Preparation

Filter aided sample preparation (FASP) was carried out using the Expedeon FASP kit (Expedeon Ltd, 44255) according to manufacturer's instructions. Briefly, 30 µl or 0.4 mg protein sample (whichever was least), previously reduced with 0.05 volumes 0.1 M DTT, was combined with 200 µl urea sample solution for protein loading on to filter membrane. The sample was loaded on to the spin filter by centrifugation at 14 000 g for 15 minutes and washed with 200 µl urea sample solution before centrifuging for a further 15 minutes at 14 000 g. Reduced disulfide bonds were alkylated by incubation with 10 µl 10 X iodoacetamide solution in 90 µl urea sample solution for 20 minutes in the dark. After removal of the iodoacetamide solution by centrifugation, samples were washed three times with urea sample solution and three times with 50 mM ammonium bicarbonate solution before incubation with 1:100 enzyme-to-protein ratio of mass spectrometry grade trypsin (Promega, V528A) for 18 hours at 37 °C. Spin filters (with bound freshly-digested peptides) were transferred to a fresh tube and peptides collected by two sequential washes with 40 µl 50 mM

ammonium bicarbonate and a further wash with 50 µl 0.5 M NaCl which will also inhibit further trypsin activity. Samples were then acidified with 1 µl 1% TFA and freeze-dried to reduce their volume before re-suspension in 3% acetonitrile, 0.1% formic acid for LC-MS/MS analysis.

2.17.4 Sample fractionation

2.17.4.1 StageTip Fractionation

StageTips were prepared and used as described by Rappsilber et al[139]. Briefly, a strong cation exchange disk was loaded into pipette tips below a C18 disk. The StageTips were primed with 20 µl methanol before sample loading and then elutions performed with varying concentrations of NH_4AcO in 0.5% formic acid, 20% acetonitrile to generate 6 fractions (20 mM, 50 mM, 100 mM, 200 mM, 400 mM, 1 M NH_4AcO). Elutions were de-salted using a C18-only StageTip prior to loading on the mass spectrometer.

2.17.4.2 Liquid chromatography

Liquid chromatography and QSTAR LC-MS/MS methods were performed as described by Smith et al.[140] and are briefly outlined below. Samples were separated on a Poly-LC strong cation exchange column (200 x 2.1 mm) at 200 µl/min by an Ettan LC (GE Healthcare) HPLC system. Peptide separation was performed using a biphasic gradient of 0-150 mM KCl over 11.25 column volumes and 150-500 mM KCl in buffer A (10 mM K_2HPO_4, 25% acetonitrile, pH 2.8) over 3.25 column volumes. A total of 30 fractions were collected and dried down before re-suspension in 2% acetonitrile/0.1% formic acid. Aliquots were then analysed by LC-MS/MS by the QSTAR system outlined below.

2.18 LC-MS/MS Analysis

2.18.1 QSTAR

Sample fractions were analysed using an ultimate 3000 UHPLC system (Dionex) attached to a hybrid quadrupole-TOF mass spectrometer (QStar Pulsar *i*, MDS-Sciex/Applied Biosystems) coupled to a nanospray source (Protana) and a PicoTip silica emitter FS360-20-10-D-20 (New Objective). Samples were loaded and washed on a Zorbax 300SB-C18,5 x 0.3 mm, 5-micron trap column (Agilent) and online chromatographic separation performed over 2 h on a PepMap™ 75 µm x 150 mm C18, 3 µm, 100 Å column (Thermo Scientific) with a linear gradient of 5-32% acetonitrile, 0.1% formic acid at a flow rate of 300 nl/minute. Analyst software version 1.1 (Applied Biosystems) was used acquire all MS and MS/MS data. Protein identifications were obtained by searching against the SwissProt 2015_04 or 2015_09 database with MASCOT version 2.5.1. Peptide mass tolerance of 50 ppm and fragment mass tolerance of 0.6 Da was used with 1 missed cleavage permitted, and variable modifications of Carbamidomethyl (C), and Oxidation (M) considered. Peptide False discovery rate was set at 1%.

2.18.2 TripleTOF 6600

2.18.2.1 Nano-flow LC-MS/MS

Sample fractions containing 1 µg peptides were analysed using an ekspert™ nanoLC 425 with nano gradient flow module (Eksigent) attached to a quadrupole Time-Of-Flight (QTOF) mass analyzer (TripleTOF 6600, SCIEX) via a nanospray III source (SCIEX) with a PicoTip emitter FS360-20-10-C12 (New Objective). Samples were loaded and washed on an Acquity M-class symmetry C18 100 Å, 180 µm x 20 mm, 5-micron trap column (Waters) and online chromatographic separation performed over 105 minutes on a Picofrit column filled with Reprosil-PUR 3 µm, 75 µm x 250 mm (New Objective) at a flow rate of 300 nl/min with a linear gradient of 30-40% acetonitrile, 0.1% formic acid over 70 minutes, then to 80% acetonitrile, 0.1% formic acid over 2 minutes, held for 6 minutes before returning to 3% acetonitrile, 0.1% formic acid and re-equilibrated. Analyst software version 1.7.1 (Applied Biosystems) was used acquire all MS and MS/MS data.

2.18.2.2 Micro-flow LC-MS/MS

Sample fractions containing 5 µg peptides were analysed using an ekspert™ nanoLC 425 with low micro gradient flow module (Eksigent) attached to a quadrupole Time-Of-Flight (QTOF) mass analyzer (TripleTOF 6600, SCIEX) coupled to a DuoSpray source (SCIEX) and a 50-micron ESI electrode (Eksigent). Samples were loaded and washed on a TriArt C18 Capillary guard column 1/32", 5 µm, 5 x 0.5 mm trap column (YMC) and online chromatographic separation performed over 57 minutes on a Triart C18 Capillary column 1/32", 12 nm, S-3 µm, 150 x 0.3 mm (YMC) at a flow rate of 5 µl/min with a linear gradient of 3-32% acetonitrile, 0.1% formic acid over 43 minutes, then

to 80% acetonitrile, 0.1% formic acid over 2 minutes, held for 3 minutes before returning to 3% acetonitrile, 0.1% formic acid and re-equilibrated. Analyst software version 1.7.1 (Applied Biosystems) was used acquire all MS and MS/MS data.

2.18.2.3 Data-Dependant acquisition

Protein identifications were obtained by searching against the SwissProt 2017_02 database in ProteinPilot™ version 5.0.1 using the Paragon search algorithm, with peptide mass tolerance 10 ppm, fragment mass tolerance 0.6 Da, 1 missed cleavage permitted and variable modification of carbamidomethyl (C) considered.

2.18.2.4 Data-Independent acquisition

Samples were spiked with iRT peptides (Biognosys) at a ratio of 1 µg protein to 0.1 µl 10 x RT peptide mix. Protein identifications were obtained by searching spectra against the panhuman 10000 protein 2014 spectral library[141] in PeakView version 2.2 with the SWATH acquisition microapp version 2.0. Chromatographic retention time calibration was performed using iRT peptides, and SWATH data processing carried out with the following filters: 300 peptides per protein, 5 transitions per peptide, 95% peptide confidence threshold, 1% false discovery rate threshold and XIC width 75 ppm. Further analysis was carried out using MarkerView version 1.2.1 before export to Microsoft office Excel and R.

2.19 Bioinformatics

The R scripts for the bioinformatic analysis described here were written by the author of this thesis and can be found in appendix 1.

Protein identifications were mapped to the gene ontology (GO) terms for biological process and cellular compartment either manually, using terms listed on UniProt, or by ClusterProfiler[142] through Bioconductor/R. Statistical testing for overrepresentation or enrichment of GO terms was performed using the Panther tools available at pantherdb.org[143,144].

Basic comparative analysis, including PCA and t-test analysis was performed in MarkerView version 1.2.1. ANOVA was performed using GraphPad Prism 7. More complex analysis and graphics generation that was not accessible by Microsoft office Excel was performed in R. Volcano plots and heat maps were created using ggplot2. Venn diagrams were created using the draw.pairwise.venn and draw.quadruple.venn functions. P-values used in quantitative analysis were adjusted using the p.adjust function (see Appendix 1: R scripts).

Chapter 3: Mass spectrometry as a tool for proteomic analysis: method development and the secretome

3 Mass spectrometry as a tool for proteomic analysis: method development and the secretome

3.1 Introduction

This chapter describes the method development lying behind much of the proteomic investigations described in later chapters. This includes the investigation into different fractionation techniques and the use of different detergents to modulate the proteomic identifications possible using tandem mass spectrometry. The QSTAR Pulsar (Applied Biosystems) machine, a hybrid quadrupole/TOF mass spectrometer, was used to collect most of the data discussed in this chapter (Figure 3.1). It contains three quadrupole units: the first (Q0) acts only as a channel to transmit and focus the ions following ionisation at the point of injection; the second (Q1) acts as a mass filter to select precursor ions of specific sizes; the third (Q2) acts as the collision cell, fragmenting the parent ions by collision induced dissociation[145].

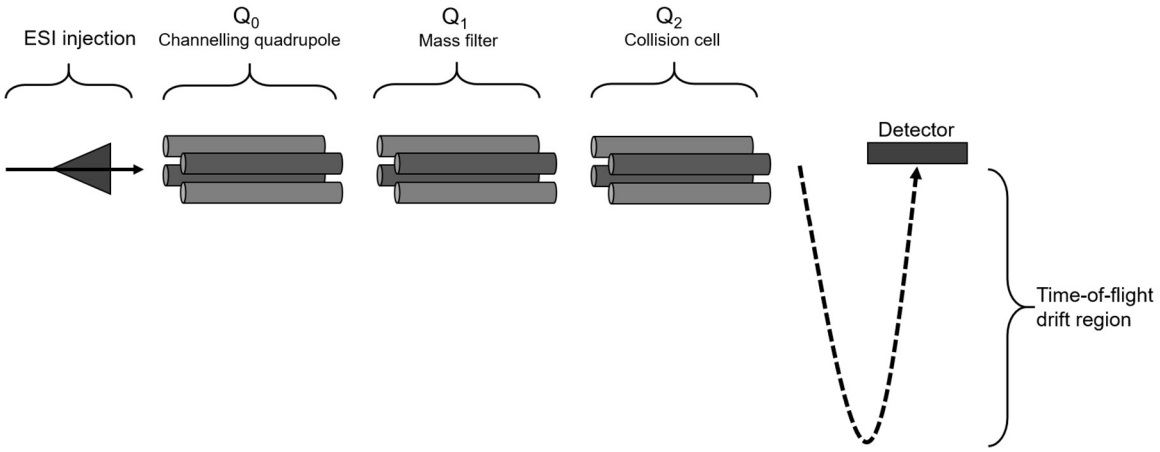


Figure 3.1: Components of QSTAR Pulsar mass spectrometer. After the ESI injection orifice, ions pass through a series of three quadrupoles: the first to channel and direct ions; the second acts as a mass filter to elect ions of certain masses; the third acts as a collision cell to fragment ions before they pass through the time of flight drift region. The time taken to travel through the time of flight tube, and the intensity of those ions, are recorded in the detector to generate mass spectra.

Also discussed in this chapter is the use of mass spectrometry for secretome analysis, which often proves to be very challenging. One of the major hurdles in secretome analysis is that cells are grown in complete media, containing 10% FCS of which one of the main constituents is BSA. This means that when analysing the mass spectrometry data, it is very hard to identify secreted proteins, which may be at relatively low concentrations in the volume of media, above a background of serum peptides. This is particularly the case with data-dependant acquisition (DDA), as used by mass spectrometers such as the QSTAR. Here, only the most abundant peptides are selected for fractionation in MS2 and therefore identification. Thus, if nearly all the most abundant peptides come from FCS, the possible identifications from secreted proteins is limited. Several methods used

to reduce FCS peptide abundance are discussed in this chapter and, in the final section, the use of data-independent acquisition (DIA) by a TripleTOF 6600 (AB SCIEX) mass spectrometer for secretome investigations is described.

The TripleTOF mass spectrometer is a more advanced machine that is similar in set up to the QSTAR but has improved ion production, capture, focussing, and detection to improve the overall sensitivity of the machine. In addition, it permits the use of DIA, which can increase the number of successful protein identifications. DIA is a method whereby, instead of selecting precursor ions for fragmentation, fragment ion spectra are generated for all precursor ions within the MS1 space[146]. The spectra are then aligned using calibration peptides, which are spiked into samples at a known and consistent concentration and have a known retention time. The remaining unknown peaks can then be identified by comparison to a spectral library which is generated using the DDA method[146]. Relative quantification is achieved by measuring summed peak areas of all fragment ions from a peptide. Two such techniques currently exist: MS^E (also known as MS^{All}) developed by Waters and sequential window acquisition of all theoretical fragment ion spectra (SWATH), which was developed by SCIEX and is used for the work described in this thesis.

3.2 LC-MS/MS sample preparation

In general, to analyse the composition of proteins in any given sample of cells, lysates were generated by disrupting cell membranes with detergents, such as Triton X-100 (TX100) or SDS, and the proteins solubilised with buffers designed to maintain protein integrity. Protease inhibitors were also used to inhibit the activity of intracellular proteolytic enzymes, which are released from the cell upon lysis. Following cell lysis, the samples were digested in-solution using trypsin to generate peptides that could be identified by mass spectrometry (MS).

A series of initial experiments were performed using HT1080 cells, derived from a human fibrosarcoma, to determine the best method for sample preparation. First, differing methods of fractionating the sample were compared before analysing the identifications found with different detergents in the lysis buffer.

3.2.1 StageTip™ fractionation of samples

StageTips™ have been identified as a method to fractionate samples on a small scale, using small discs of SCX material placed within pipette tips[139]. To assess the fractionation efficiency of StageTips™, a flask of HT1080 cells, grown to 80% confluency in MEM with 5% BSA, was lysed in 1 x MNT lysis buffer with 1% v/v Triton X-100 and protease inhibitors. The sample was digested in solution with trypsin, and some set aside for analysis without fractionation. The remaining sample was then fractionated with strong cation exchange StageTips™. StageTips™ are designed to work like a small-scale chromatography column, allowing the fractionation of peptides according to their charge, through a pipette tip[139]. Samples were fractionated into six fractions, with peptides eluted with ammonium acetate at concentrations of 20 mM, 50 mM, 100 mM, 200 mM, 400 mM, and 1 M. After the elution, the ammonium acetate salts require removal before MS, therefore each fraction was put through another StageTip™, this time with a C18 desalting disk. All fractions, as well as the unfractionated sample, were then analysed in the mass spectrometer and the number of identifications compared between the unfractionated and fractionated sample. Many more identifications were possible with fractionation (Figure 3.2). Identifications were deemed significant when more than one peptide was matched and the false discovery rate (FDR) was less than 5%. The FDR is calculated using a decoy database where the sequences in the database are reversed and results searched against this. Any true positive matches ought to only match in the target database, and not the decoy database. The FDR is the total number of false positives (i.e. those which also matched against the decoy database), divided by the sum of all matches against the target database (these will be both true positives and false positives).

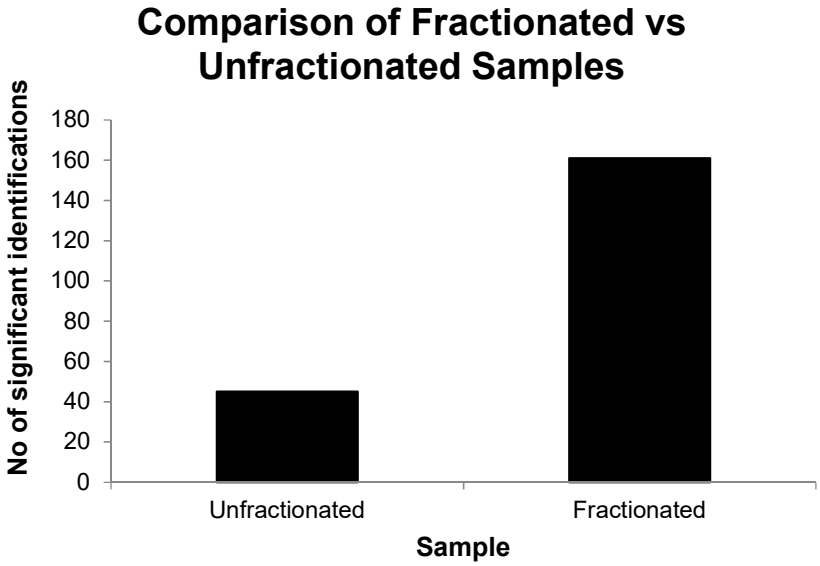


Figure 3.2: Chart depicting the difference in the number of significant proteins identified in samples that were fractionated or left unfractionated prior to MS analysis. Significant identifications were determined as those with more than 1 peptide match per protein, with a false discovery rate of <5%.

3.2.2 StageTip™ vs strong cation exchange column fractionation

Samples may also be fractionated using a strong cation exchange column on a liquid chromatography machine. This generates many more fractions, across a longer gradient, which may increase protein identifications further than the numbers seen with StageTip™ fractionation. Whereas with StageTips™ 6 fractions were collected, the liquid chromatography column generated 30 fractions along an elution gradient. The total number of identifications from each fractionation technique, and the number of significant identifications were compared (Figure 3.3). Whilst there were nearly 100 more total proteins identified with column fractionation, there were only four more significant IDs. Although maximising the number of identifications is important, the increase in time both in person hours, and time on the MS machine to run many more fractions, was deemed to be too great for such a minor increase in significant identification numbers.

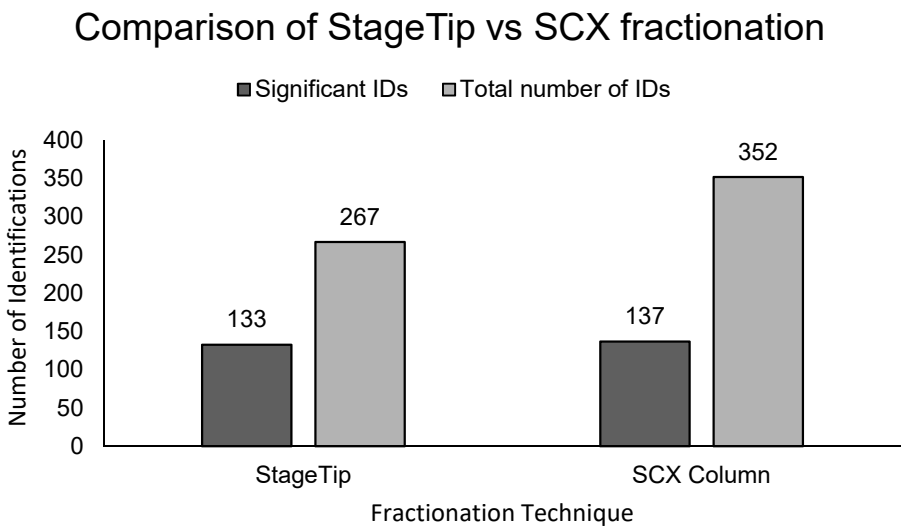


Figure 3.3: Chart depicting the difference in the number of proteins identified with different fractionation techniques. Total number of IDs (light grey) includes both significant and insignificant results. Significant IDs (dark grey) are just those with more than one peptide match, and FDR<5%. Comparison between strong cation exchange (SCX) column fractionation, and StageTip™ fractionation.

3.2.3 Detergent comparison of cell lysis

TX100 is a typical detergent used for mammalian cell lysis, however this detergent cannot, on its own, be used in mass spectrometry as the PEG components lead to a repeating series of high intensity peaks, 44 Da apart on the spectrum[147]. In data-dependant mass spectrometry, like that used in these experiments for protein identification, only the peaks with the highest intensities are selected for MS2 analysis, and therefore the presence of PEG peaks leads to the exclusion of many, if not all, peaks from legitimate peptide precursor ions. In order to avoid this issue, when preparing protein samples from cell lysates produced using TX100, cell lysis was followed with an acetone precipitation step and the resulting protein pellets were then resuspended in 50 mM ammonium bicarbonate (ABC) with 2% SDS to aid re-solubilisation[148]. A range of other detergent conditions were also trialled to investigate the effect on protein identifications. These conditions were: Lauryl Maltose Neopentyl Glycol (LMNG) with and without acetone precipitation, and sodium dodecyl sulfate (SDS).

LMNG is a member of the Maltose–neopentyl glycol (MNG) family of amphiphiles, which were developed to aid the solubilisation and crystallisation of integral membrane proteins[149]. MNGs were determined to be better than conventional detergents at conserving native protein structure, and also have a very low critical micelle concentration (cmc) value[149]. This means that the number of detergent molecules required to form micelles[150] around the proteins to be solubilised is much lower than with conventional detergents whose cmc value may be much higher. This is beneficial for mass spectrometry because it means lower concentrations of detergents can be used, therefore reducing the potential interference with the machine. LMNG was used to investigate whether it aided the recovery of a subset of proteins that were not seen with other detergents.

SDS is a strong anionic detergent that rapidly lyses cells, but will also partially denature proteins[151]. This provides good solubilisation of proteins; however, it does not allow efficient digestion by trypsin in the next stage of sample preparation. Analysis by LC-MS/MS can generally tolerate SDS concentrations of around 0.01%[152], however this is considerably lower than the concentrations typically used in cell lysis (around 10%). Therefore, when lysing samples in SDS, acetone precipitation was also performed as for TX100 samples, and resuspended in ammonium bicarbonate with a lower SDS concentration, in a volume that could be diluted to tolerable levels.

To assess if there was a difference in the profile of proteins identified with different detergents, a series of HT1080 cell lysates were generated using the different detergent conditions outlined above (section 3.2), fractionated using StageTip™ fractionation (section 3.2.1), and the identifications from each fraction combined into a single list. GO terms were then assigned to each protein by using the GO terms listed on the UniProt database and assigning the term that most wholly described the proteins function to a list of protein identifications. The distribution of these

terms is represented graphically in Figure 3.4. Comparison of the graphs produced demonstrated that there was no obvious difference in the set of proteins identified with the different detergents. That is that, across the different detergent conditions, a similarly broad range of biological processes and balance of cellular compartments was seen with no condition leading to the identification of more proteins annotated to a specific ontology term. With the SDS detergent, more nuclear proteins were identified (5% of identifications compared to 3 or 4% in other conditions). However, this is expected because SDS can lyse the nuclear membrane.

This experiment confirmed that the acetone precipitation of lysates prepared with TX100, prior to digestion, was a suitable method that led to sufficient protein identification without unexpected bias towards subsets of proteins at the location or function level. In addition, lysis with TX100 appeared to identify a higher proportion of ER associated proteins compared to other detergent conditions, which will be an advantage in the investigation in to the effect of ER stress. SDS would be an alternative to use in order to gain recovery of nuclear proteins, while LMNG does not offer an obvious advantage in terms of protein recovery for these types of experiments.

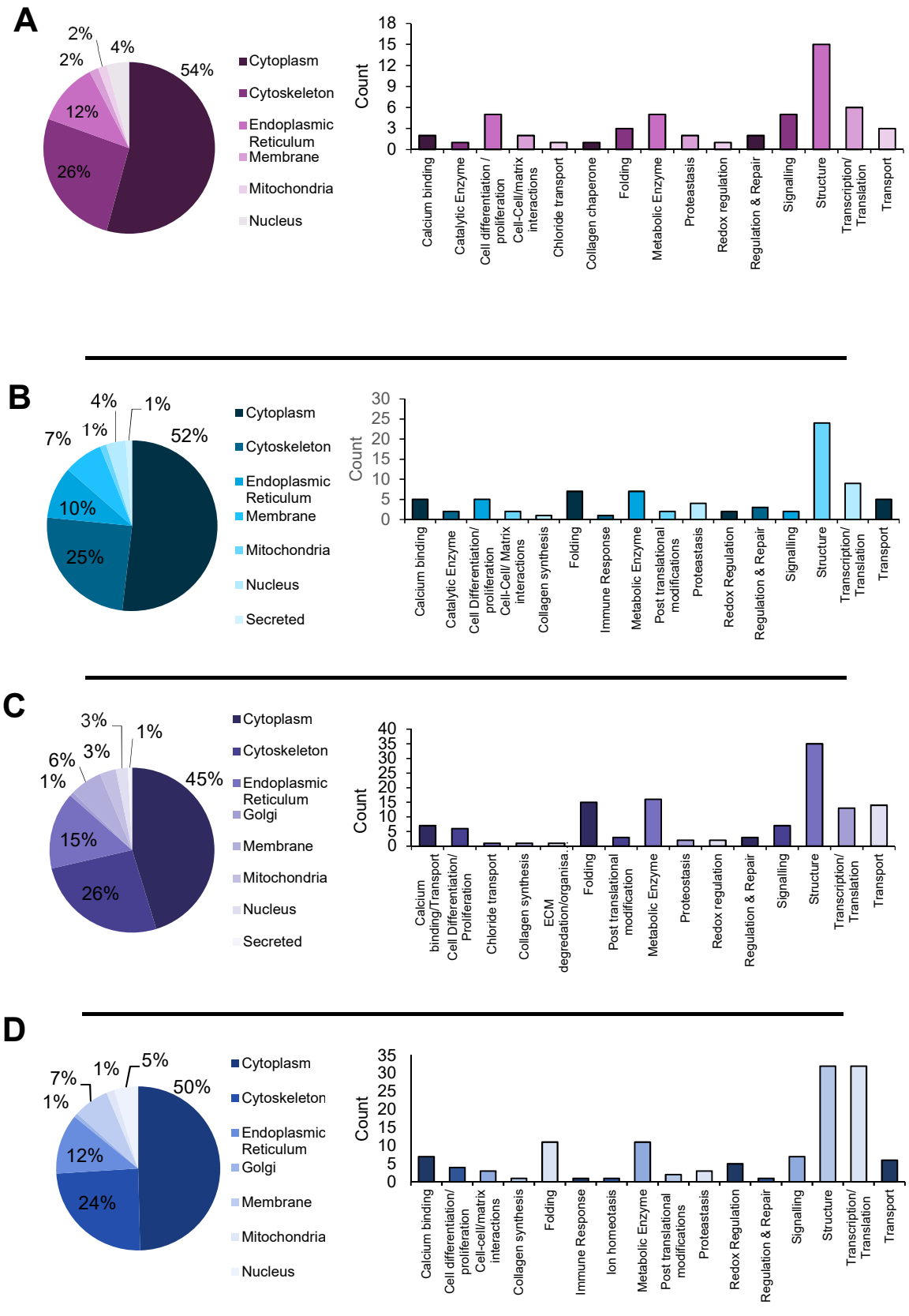


Figure 3.4: Subcellular distribution, and biological functions of proteins identified from lysates with different detergents. Proteins manually annotated with GO terms from Uniprot. Pie chart (left) depicts subcellular location distribution while bar chart (right) shows function distributions. Different detergent conditions shown in each panel were as follows: **A.** LMNG with acetone precipitation. **B.** LMNG alone **C.** Triton X -100. **D.** SDS.

3.3 Sensitivity and Reproducibility analysis

3.3.1 Immunoprecipitation

While investigating the capabilities of the QSTAR machine, and optimising protocols, immunoprecipitation was used as a technique to test the MS sensitivity. The specificity of an immunoprecipitation means that there will be fewer proteins, and therefore fewer peptides in the sample. A more sensitive machine should be able to identify more peptides under these conditions. In addition, this technique is a useful tool to potentially identify unknown interactions between proteins. HT1080 cells were lysed in TX100 lysis buffer, and the lysate then immunoprecipitated with an antibody to Ero1α. This was designed to isolate Ero1α, along with any interacting proteins, from the cell lysate. Ero1α was used as the target in these experiments as its interaction with PDI is well known and therefore could act to validate the results. These proteins reside in the ER, and are required for the correct formation of disulfide bonds during the folding of proteins in the early secretory pathway[153] (see section 1.2.2.4).

Typically, immunoprecipitation requires several stringent washing steps, which must include detergent, in order to remove proteins which are bound to the target protein by non-specific interactions[154]. Unfortunately, these detergents then lead to problems downstream at the LC-MS/MS stage (in the same way as discussed in reference to TX100 above, section 3.2.3). Therefore, the initial protocol was modified to alleviate the problems caused by the necessity for detergent in the washing steps.

Firstly, samples were washed twice with the same buffer as that used for cell lysis namely MNT buffer with 1% TX100, before four further washes with MNT buffer containing no detergent. Next, the proteins were eluted in a buffer containing 1% SDS and DTT, instead of high levels of detergent. The eluted samples could then be diluted and directly digested with trypsin and identifications made by LC-MS/MS, without detergent interference.

The protein concentration of the eluates following immunoprecipitation with Ero1α was estimated by Bradford assay to be much less than 1 µg/µl. A standard injection to the mass spectrometer is 1 µl, so this technique therefore tested the sensitivity of the machine as any proteins identified above a background of antibody and protein A were inevitably present at low levels.

From these immunoprecipitation experiments, the target protein (Ero1α), and protein A which is bound to the sepharose beads, were confidently identified with more than one peptide match (Table 3.1; Figure 3.5). Unfortunately, however, it was not possible to identify other interacting proteins. This suggests that there were limitations to the either the protocol or sensitivity of this machine, and it was not possible to confidently identify co-immunoprecipitating proteins that were not abundant in the samples.

Table 3.1: Identifications from Immunoprecipitation with Ero1. Identifications from mass spectrometry analysis of immunoprecipitation eluent were protein A (SPA_STAA8), and the target protein Ero1 (ERO1A_HUMAN).

Accession	Mass	No. of sequences	No. of significant sequences	Description
SPA_STAA8	56403	5	4	Immunoglobulin G-binding protein A OS=Staphylococcus aureus (strain NCTC 8325) GN=spa PE=1 SV=3
ERO1A_HUMAN	54358	2	2	ERO1-like protein alpha OS=Homo sapiens GN=ERO1L PE=1 SV=2

A

1	MKKKNIYSIR	KLGVGIASVT	LGTLLISGGV	TPAANAAQHD	EAQQNAFYQV
51	LNMPNLNADQ	RNGFIQSLKD	DPSQSANVLG	EAQKLNDSQA	PKADAQQNNF
101	NKDQQSAFYE	ILNMPNLNEA	QRNGFIQSLK	DDPSQSTNVL	GEAKKLNESQ
151	APKADNNFNK	EQQNAFYEIL	NMPNLNEEQR	NGFIQSLKDD	PSQSANLLSE
201	AKKLNESQAP	KADNKFNKEQ	QNAFYEILHL	PNLNEEQRNG	FIQSLKDDPS
251	QSANLLAEAK	KLNDAQAPKA	DNKFNKEQQN	AFYEILHLPN	LTEEQRNGFI
301	QSLKDDPSVS	KEILAEAKKL	NDAQAPKEED	NNKPGKEDNN	KPGKEDNNKP
351	GKEDNNKPGK	EDNNKPGKED	GNKPGKEDNK	KPGKEDGNKP	GKEDNKKPGK
401	EDGNKPGKED	GNKPGKEDGN	GVHVVKPGDT	VNDIAKANGT	TADKIAADNK
451	LADKNMIKPG	QELVVDKKQP	ANHADANKAQ	ALPETGEENP	FIGTTVFGGL
501	SLALGAALLA	GRRREL			

B

1	MGRGWGFLFG	LLGAVWLLSS	GHGEEQPPET	AAQRCFCQVS	GYLDDCTCDV
51	ETIDRFNNYR	LFPRLQKLLE	SDYFRYYKVN	LKRPCPFWND	ISQCGRRDCA
101	VKPCQSDEVP	DGIKSASYKY	SEEANNLIEE	CEQAERLGAV	DESLSEETQK
151	AVLQWTKHDD	SSDNFCEADD	IQSPEAEYVD	LLLNPERYTG	YKGPDAWKIW
201	NVIYEENCFK	PQTIKRPLNP	LASGQGTSEE	NTFYSWLEGL	CVEKRAFYRL
251	ISGLHASINV	HLSARYLLQE	TWLEKKWGHN	ITEFQQRFDG	ILTEGEGPRR
301	LKNLYFLYLI	ELRALSKVLP	FFERPDFQLF	TGNKIQDEEN	KMLLLEILHE
351	IKSFPLHFDE	NSFFAGDKKE	AHKLKEDFRL	HFRNISRIMD	CVGCFKCRLW
401	GKLQTQGLGT	ALKILFSEKL	IANMPESGPS	YEFHLTRQEI	VSLFNAFGRI
451	STSVKELENF	RNLLQNIH			

Figure 3.5: Amino acid sequences of identified proteins of interest. Peptides which were successfully identified are in red. **A.** Protein A with 5 peptides identified. **B.** Ero1A with 2 peptides identified.

Another way to identify protein interactions following immunoprecipitation is to separate the eluent on a 2D gel, select the spots, digest the proteins from within the gel and analyse these by mass spectrometry. To test this approach, and the possibility of using a 2D method to identify novel interactions, a liver tissue sample was lysed and immunoprecipitated with an antibody raised against PDI. Liver tissue provided a large, easily obtainable sample that is rich in protein.

Three 2D gels were run: one each for mock PDI antibody, liver tissue lysate, and the immunoprecipitation eluate (Figure 3.6). From the eluate gel, three spots were removed, and proteins digested for identification by mass spectrometry, which were initially assumed to be PDI, antibody heavy chain and antibody light chain.

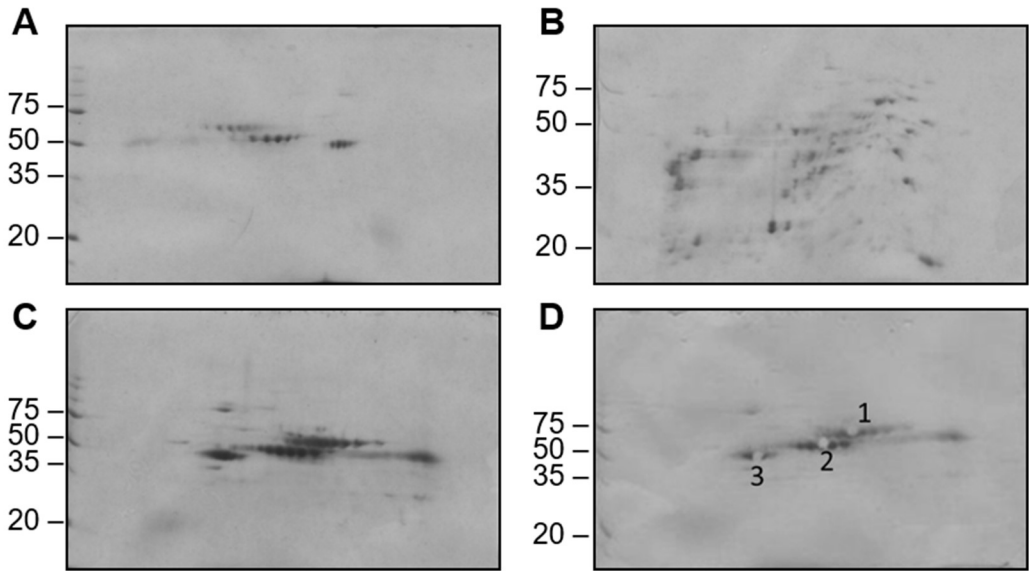


Figure 3.6: 2D gel analysis of liver tissue. Proteins were separated first on a pH 4-11 IEF strip, then on 10% SDS PAGE gels and stained with colloidal Coomassie for the visualisation of protein spots. Inputs were: **A.** PDI antibody only; **B.** liver tissue lysate only; **C.** liver tissue immunoprecipitated with anti-PDI. **D.** 2D Gel Post spot picking to display location of spots picked from eluate gel that were digested for protein identification by mass spectrometry.

Following analysis, PDI was not identified in any of the spots; the only expected protein was IgG gamma chain. Interestingly, fibrinogen beta chain was also identified in spot 1 and fibrinogen alpha chain was detected in spot 2 (Table 3.2). Fibrinogen is a glycoprotein that circulates in the blood with prothrombin. At a wound site, factor VIII aids the conversion of prothrombin to thrombin which then catalyses the proteolysis of fibrinogen into fibrin monomers which form a blood clot[155]. Although a direct interaction between PDI and fibrinogen has not yet been identified, several recent publications have established a role for PDI in blood clot formation [156–159] indicating that this could be a true, and novel interaction. Specifically, Swiatkowska et al., demonstrated that Ero1α is present on the platelet surface with PDI, and blocking Ero1α with specific antibodies abolished platelet responses including fibrinogen binding. Immunoprecipitations from their studies also demonstrated an interaction between Ero1α and fibrinogen[159]. Further investigations would be required to determine whether the interaction between PDI and fibrinogen seen in our work is direct or indirect, perhaps involving Ero1α.

In addition to detecting a novel PDI interacting protein, many expected contaminating keratin proteins were identified. Trypsin was also identified as a remnant from the digestion process. Several species isoforms were also detected for proteins. Where there is homology in the protein sequences of a given protein between species, it is possible for peptides to be identified as the same protein in more than one species. For example, in spot 3 argininosuccinate synthase was identified with 14 peptides as mouse and 7 peptides as human (Table 3.2). This is a common occurrence in database searches where taxonomy has not been specified, as was the case in this database search to permit identification of rabbit IgG, which confirms successful elution.

Table 3.2: Mass spectrometry identifications from proteins in spots taken from 2D Gel of PDI immunoprecipitation eluate. Spot numbers correspond to those in Figure 3.6D. Accession number, number of peptide sequences identified, number of significant peptide sequences identified, and descriptions are given for each protein identification. Significant sequences are those that are deemed to have met a scoring threshold based on the number of y and b ions in a peptide spectral match, as well as the number of peptide spectral matches per sequence. Proteins referred to in the text are highlighted in green.

Spot No.	Accession	No. of sequences	No. of significant sequences	Description
1	K2C1_HUMAN	9	9	Keratin, type II cytoskeletal 1
	K22E_HUMAN	9	9	Keratin, type II cytoskeletal 2 epidermal
	FIBB_MOUSE	10	10	Fibrinogen beta chain
	FIBB_HUMAN	8	8	Fibrinogen beta chain
	FIBB_CHICK	3	3	Fibrinogen beta chain (Fragment)
	K1C10_HUMAN	10	10	Keratin, type I cytoskeletal 10
	K1C9_HUMAN	6	6	Keratin, type I cytoskeletal 9
	FIBB_RABIT	3	3	Fibrinogen beta chain (Fragment)
	TRYP_PIG	2	2	Trypsin
	CASA1_BOVIN	3	3	Alpha-S1-casein
	HORN_HUMAN	2	2	Hornerin
	IGHG_RABIT	3	3	Ig gamma chain C region
	CASB_BOVIN	2	2	Beta-casein
2	K2C1_HUMAN	9	9	Keratin, type II cytoskeletal 1
	K1C10_HUMAN	6	6	Keratin, type I cytoskeletal 10
	TRYP_PIG	2	2	Trypsin
	IGHG_RABIT	5	5	Ig gamma chain C region
	FIBA_MOUSE	3	3	Fibrinogen alpha chain
	FIBA_HUMAN	3	3	Fibrinogen alpha chain
3	K22E_HUMAN	18	18	Keratin, type II cytoskeletal 2 epidermal
	K2C1_HUMAN	15	15	Keratin, type II cytoskeletal 1
	K2C5_HUMAN	5	5	Keratin, type II cytoskeletal 5
	K1C10_HUMAN	16	16	Keratin, type I cytoskeletal 10
	ASSY_MOUSE	14	14	Argininosuccinate synthase
	ASSY_HUMAN	7	7	Argininosuccinate synthase
	K1C9_HUMAN	10	10	Keratin, type I cytoskeletal 9
	TRYP_PIG	2	2	Trypsin
	THIM_MOUSE	2	2	3-ketoacyl-CoA thiolase, mitochondrial
	K1C17_MOUSE	2	2	Keratin, type I cytoskeletal 17
	CPSM_MOUSE	2	2	Carbamoyl-phosphate synthase [ammonia], mitochondrial

3.3.2 Interaction networks to assess reproducibility

Bioinformatics tools can be used to assess protein-protein interactions. STRING.db, for example, [160] is an online tool, into which protein identification(s) can be inputted and it will suggest interacting proteins, based on a range of peer-reviewed sources. SERPINH1, otherwise known as heat shock protein 47 (HSP47) was a protein identified from the proteomic analysis of whole cell lysates described previously (section 3.2). SERPINH1 is an ER resident chaperone that aids in the formation of mature collagen fibrils. It is thought to bind procollagen fibrillar domains, in order to provide thermal stability in their immature state prior to mature collagen fibril assembly[77]. Using the STRING.db programme, several proteins thought to interact with SERPINH1 were identified (Figure 3.7). These included the beta subunit of prolyl 4 hydroxylase (P4HB), which is PDI, and other ER resident chaperone proteins (see Table 3.3) which likely play roles in the correct folding of collagens, or other secretory proteins found within the ER lumen.

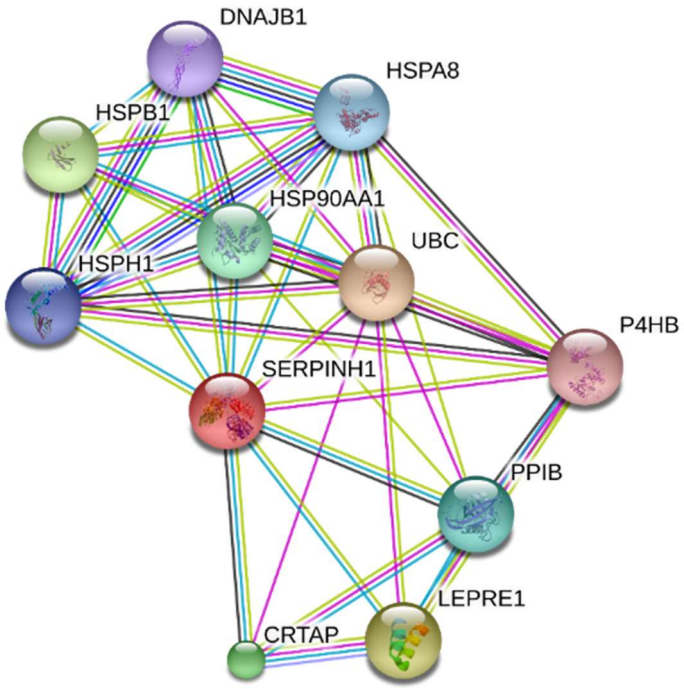


Figure 3.7: STRING-db.org suggested interaction map of SERPINH1. Coloured lines represent different types of known or predicted interactions as follows: **Turquoise** demonstrates a known interaction from curated databases; **pink** demonstrates a known interaction that has been experimentally determined; **green** demonstrates a predicted interaction based on the gene neighbourhood; **red** demonstrates a predicted interaction based on gene fusions; **blue** demonstrated predicted interaction based on gene co-occurrence; **lime** demonstrates a suggested interaction based on text mining evidence; **black** demonstrates a suggested interaction based on co-expression and finally **lilac** demonstrates a suggested interaction based on protein homology.

To assess the reproducibility of the results from the QSTAR mass spectrometer, these network identifications were used. The protocols established for cell lysis and sample digestion for proteomic analysis were transferred to nHDF cells, and an experiment was performed to establish how many proteins in the SERPINH1 network shown in Figure 3.7 could be identified in fibroblast cell lysates. Three biological replicate samples of nHDF cells were grown to confluence and lysed in MNT lysis buffer containing 1% TX100. The lysates were digested in solution and fractionated using StageTips™ before protein identification by mass spectrometry. Results from all fractions of a given sample were combined to give one data set for each biological replicate. By comparing how many of the proteins in the network were identified in three repeat samples (Table 3.3), it was possible to gain some insight into reproducibility. This demonstrated that, of the proteins identified in the network, 5 of them were identified in three repeat samples, and none were detected only once, indicating that reproducibility could be expected from experimental results obtained on this machine.

Table 3.3: Reproducibility analysis of SERPINH1 STRING-db network. Data sets from protein identifications from nHDF cell lysates in three biological replicates (samples 1 – 3) were mined for the identification of proteins found in the SERPINH1 network on String-db. Successful identifications are denoted with a tick, while unsuccessful identifications are greyed out.

Sample	SERPINH1	HSP90AA1	HSPH1	HSPB1	DNAJB1	HSPA8	UBC	P4HB	PPIB	LEPRE1	CRTAP
1	✓	✓		✓		✓		✓	✓		
2	✓	✓		✓		✓		✓	✓		
3	✓			✓		✓		✓	✓		

3.4 Secretome analysis

Collagen is also of interest when analysing fibroblast secretory function in the dermis. Fibroblasts are responsible for the provision of the extracellular matrix proteins to support the surrounding tissue. Various proteins and chemokines are released into the cells environment to enable interactions and communication with other cell types. For example, stromal cell-derived factor-1 (SDF-1) is secreted by fibroblasts in the dermis to stimulate epidermal keratinocyte proliferation[161]. ECM proteins are secreted to support the function of the skin as described in chapter 1. Therefore, when assessing the impact of different stressors on dermal fibroblast cells, it would be interesting to look at the secretory profile of these cells, as well as the intracellular proteome.

As described in the introduction to this chapter, one of the major hurdles in secretome analysis is that cells are grown in complete media, containing 10% FCS of which one of the main constituents is BSA. A series of different methods were used to assess the possibility of overcoming the limitations imposed by serum in the culture medium.

3.4.1 Concanavalin A affinity purification

Firstly, concanavalin A (conA), was used to affinity purify N-linked glycoproteins. Proteins are only N-glycosylated when they enter the secretory pathway, therefore the only proteins eluting from a conA pull-down should be secreted N-linked (unmodified) glycoproteins, or secretory pathway resident glycoproteins that have been released into the media by dead cells. Samples of conditioned media were collected from nHDF cells, grown to confluency in media containing 10% FCS, and subjected to conA affinity purification. The conditioned media, affinity purification supernatant, and the eluent were then run on a 1D SDS-PAGE gel (Figure 3.8). A large band was visible at around 50 kDa which corresponds to BSA. This band was not present in the eluent. In the eluent lane, there were some less intense bands in the higher molecular weight region of the gel, which appeared to be absent, or less abundant in the other two lanes. This suggests that this method successfully concentrated secreted proteins and reduced the BSA concentration. The ~25 kD band which is present only in the eluent lane is thought to be conA itself, released from the beads during elution.

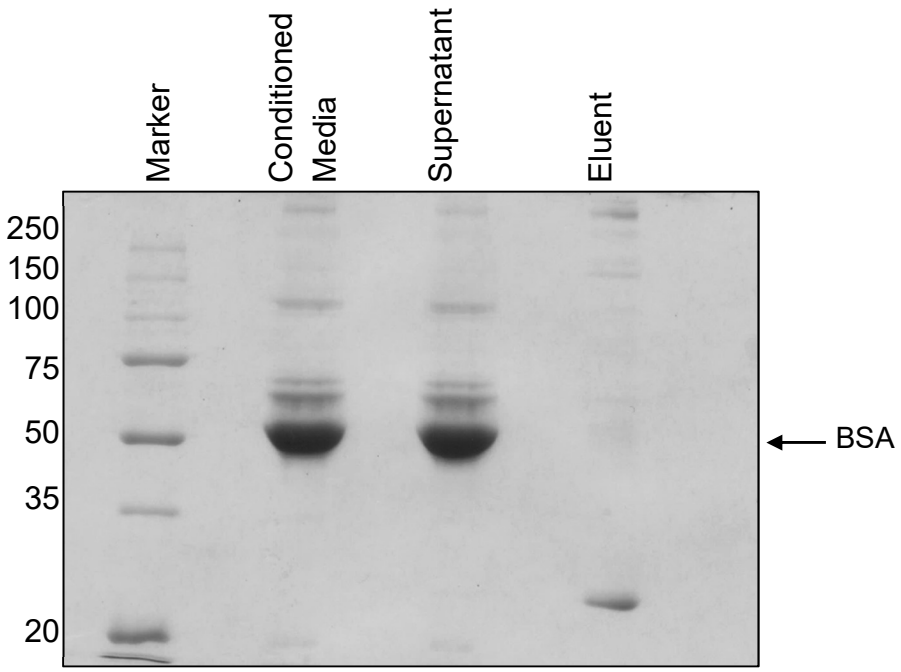


Figure 3.8: SDS-PAGE gel of conditioned media, supernatant and eluate from concanavalin A affinity purification. Stained with colloidal Coomassie stain. BSA is shown, present in both conditioned media and supernatant from the affinity purification.

To test whether this method could be used for the identification of secreted proteins, the conA purification was repeated with three biological replicates of conditioned media from nHDF cultures. The eluents were again run on a 10% SDS-PAGE gel and six bands from the higher molecular weight region of gel were excised for protein identification by mass spectrometry (Figure 3.9). These were identified as bovine complement C3 and alpha -2-macroglobulin (α2M) (Table 3.4). This protein is also found in FCS and may therefore be a bovine serum contaminant. Although conA does not bind to the BSA, other serum proteins remained in the ConA eluted samples and may give false positives in the identification of proteins from conditioned media samples. A media only control in this experiment may aid in the identification of proteins unique only to conditioned media eluents.

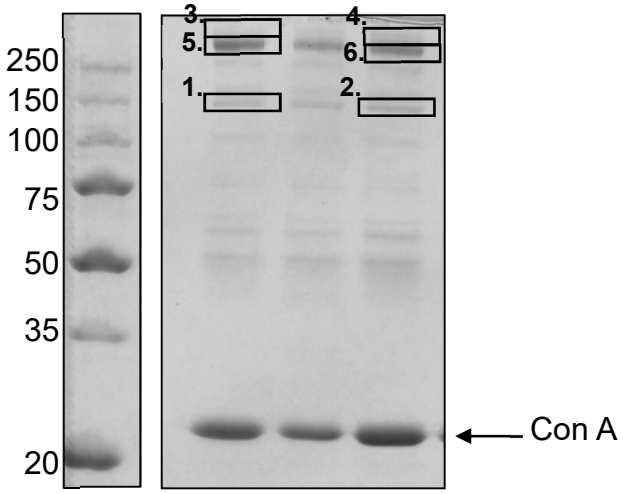


Figure 3.9: SDS PAGE gel of conA eluents, indicating location of excised bands for mass spectrometry analysis. Eluents from the conA affinity purification of three biological replicates of conditioned media from nHDF cultures were run on a 10% SDS PAGE gel. The bands which were excised for protein identification by mass spectrometry are highlighted with a box and numbered. The numbers correspond to table 4 below.

Table 3.4: Protein identification from excised gel bands displayed in Figure 3.9. Protein identifications from the mass spectrometry analysis of peptides digested from excised gel bands numbered 1 – 6. All proteins identified were contaminants from bovine serum.

Band	Accession	No. of sequences	No. of significant sequences	Description
1	CO3_BOVIN	70	56	Complement C3 OS=Bos taurus GN=C3 PE=1 SV=2
2	CO3_BOVIN	69	39	Complement C3 OS=Bos taurus GN=C3 PE=1 SV=2
3	A2MG_BOVIN	36	3	Alpha-2-macroglobulin OS=Bos taurus GN=A2M PE=1 SV=2
4	A2MG_BOVIN	53	39	Alpha-2-macroglobulin OS=Bos taurus GN=A2M PE=1 SV=2
5	A2MG_BOVIN	63	46	Alpha-2-macroglobulin OS=Bos taurus GN=A2M PE=1 SV=2
6	A2MG_BOVIN	67	59	Alpha-2-macroglobulin OS=Bos taurus GN=A2M PE=1 SV=2

3.4.2 Molecular weight cut off spin columns

Next, ways to increase the concentration of secreted proteins were investigated to increase the likelihood of detecting secreted proteins in the relatively low volumes used for mass spectrometry. Spin columns with a 3000-molecular weight cut off were used to concentrate proteins in the conditioned media. Liquid and any molecules with a molecular weight of less than 3000 can pass through the filter, leaving proteins in a lower volume on top of the filter. The concentrated samples can be removed from the column and analysed.

Samples of conditioned media from nHDF cells, both before and after concentration in a spin column, were subject to a Bradford assay to determine protein concentrations (Table 3.5) and then proteins separated on 1D SDS-PAGE gels (Figure 3.10). It was clear that the spin columns greatly increased the concentration of protein detectable in the conditioned media. However, as was seen by comparing the band patterns on the two gel images, the serum proteins were also greatly concentrated and therefore the QSTAR mass spectrometer was still unable to detect secreted proteins above the background of serum. It was concluded that to facilitate secretome analysis in the future, it will likely be necessary to grow cells in serum free conditions, develop methods to specifically deplete serum or use a more sensitive mass spectrometer.

Table 3.5: Estimation of protein concentrations of media samples. A Bradford assay was used to determine the concentration of protein in conditioned media before and after concentration using 3000 MWCO tubes, and the flow through from these tubes.

Sample	Absorbance	Protein Concentration ($\mu g\ \mu l^{-1}$)
Media	0.433	<<1
Concentrated media	0.984	2.46
Flow through	0.022	<<1

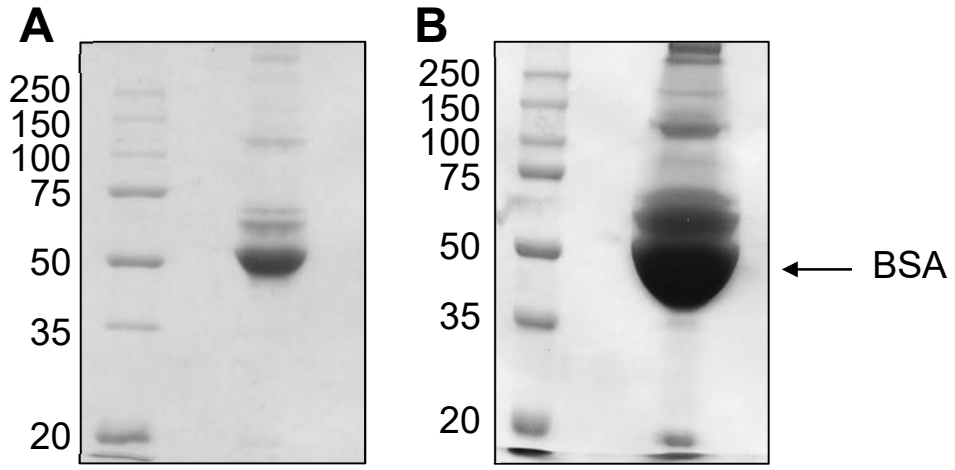


Figure 3.10: Comparison of proteins following use of a molecular weight cut off spin column. SDS-PAGE gels were stained with colloidal Coomassie. **A.** SDS-PAGE gel of conditioned media before concentration. **B.** SDS-PAGE gel of conditioned media after concentration.

3.5 Secretome analysis with a TripleTOF 6600

The investigations so far have relied heavily upon shotgun proteomics approaches, where precursor ions are selected from MS1 spectra for fragmentation and fragment ion spectra are then matched against a database for peptide identification, termed data-dependant acquisition (DDA). Using more powerful mass spectrometers, with better sensitivity and improved chromatography, increases the number of identifications, and permits quantification of peptides by more complex mass spectrometric techniques.

This improvement in mass spectrometry capability facilitated re-analysis of the fibroblast secretome previously discussed (section 3.4). It also permitted the relative quantitation of results, enabling the analysis of data not only by identifying proteins, but also by comparing the relative abundances of proteins between different treatment conditions. Having established analytical methods in HDF cells, BJ fibroblasts were used for subsequent experiments. These are a commercially available, telomerase negative cell line with a longer lifespan than other human dermal fibroblast cell lines, meaning they can be cultured for more population doublings before senescence occurs[162]. They are also well characterised by the industrial partners to this project, Procter & Gamble (P&G).

To harness the quantitative capability of the TripleTOF mass spectrometer, cells were treated with either DTT, etoposide or control medium. DTT is a reducing agent that induces an ER stress response within cells and will be used in subsequent experiments in this thesis to elucidate proteomic responses to reductive stress. Etoposide is a chemical that induces DNA damage within the cells and drives them into senescence. Senescent cells are neither apoptotic, nor active, and they are no longer able to divide or grow. They also possess a well-defined senescence-associated secretory phenotype (SASP)[163]. BJ fibroblasts were grown to 70% confluency were treated for 6 hours with 5 mM DTT, 12.5 µM etoposide, or left untreated (mock) before being given fresh media. After 4 days of recovery, the spent media was collected for analysis. Etoposide served as a positive control to determine whether 4 days in the same media drove cells to senescence without treatment. There is no knowledge in the literature that suggests reductive stress leads cells to senescence, and therefore it was of interest to determine whether DTT treated cells shared any characteristics with etoposide treated cells.

The conditioned media was collected and prepared for mass spectrometry analysis. As the TripleTOF machine was much more sensitive to detergents, digests were carried out using FASP. This sample preparation technique was described by Wiśniewski et. al., as a method that successfully depletes detergent from a sample by buffer exchange into urea[164]. Essentially proteins are bound onto an ultrafiltration device, washed and concentrated into a urea buffer. Whilst on the filter, reduction, alkylation and digestion occur, after which peptides can be spun through the

filter into a collection tube and the digestion stopped by acidification (see section 2.17.3). This was the standard sample preparation method used with the TripleTOF mass spectrometer, and so this method was also employed for the preparation of conditioned media for comparability.

A single biological replicate was used with three technical replicates of the SWATH acquisition. Using this method, 776 proteins were successfully identified, and their relative intensities quantified in all three treatment categories: untreated, DTT, and etoposide.

Looking firstly at only the identifications from the DIA data, without consideration for the quantified values, it was obvious that many more secreted proteins were identified than in the DDA data. GO annotations were assigned to the identifications using the groupGO function of ClusterProfiler in R[165]. The GO annotation for cellular compartment reveals that approximately 50% of proteins were annotated to extracellular regions, with other frequently observed annotations including membrane and cell-junction proteins (Figure 3.11). There were fewer terms associated with intracellular compartments, indicating that many of the proteins identified were "true" secretome identifications. It is important to remember here that, using this method for GO annotation, proteins are mapped to several GO terms, which could cause substantial overestimation of the numbers of proteins identified that are found in particular cellular compartments. For example, ER proteins such as PDI are annotated both with membrane-enclosed lumen [ER] and extracellular region, due to their association with the secretory pathway.

The identifications were therefore also tested for statistical overrepresentation of GO terms, compared to the human genome. Using Pantherdb.org tool for statistical overrepresentation test, the list of identifications was tested against a reference list of all known human genome proteins using the Panther GO-Slim Cellular component data set for annotation, and the Fisher's Exact with FDR multiple test correction test[143,166]. This determined statistically significant overrepresentation of several terms, including extracellular matrix (fold enrichment 2.56, p=6.83×10^{-4}) and extracellular space (fold enrichment 1.66, p=1.41×10^{-3}) (Table 3.6). The ECM proteins identified included collagens COL1A2, and COL3A1, as well as neutrophil collagenase matrix metalloproteinase 8 (MMP8), and metalloproteinase inhibitor tissue inhibitor of metalloproteinases 1 (TIMP1), which may both be associated with regulation of ECM re-modelling. All the ECM proteins associated with this overrepresentation are listed in Table 3.7. Whilst collagen proteins are important in conferring structural integrity to the ECM, the degradation of collagen fibres is tightly regulated to modulate the freedom with which cells such as fibroblasts can migrate through the matrix. For example, greater motility is required during wound healing so that fibroblasts can reach the wound site and deposit new ECM. The quantitative data from this experiment will determine whether these proteins are regulated by specific treatments or are part of a more general secretome of these cells.

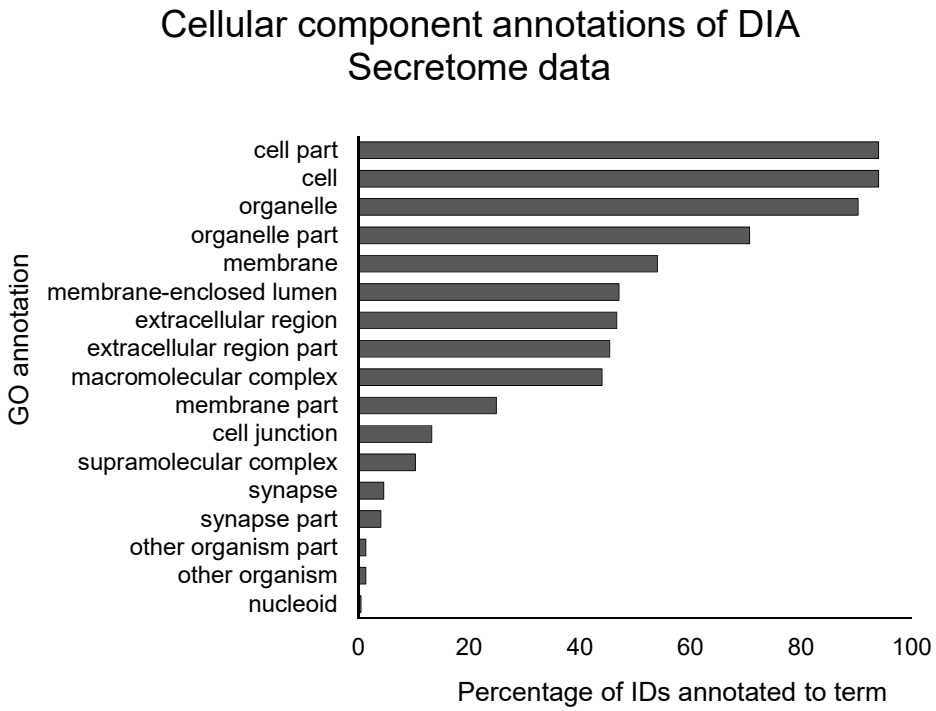


Figure 3.11: 46% of proteins identified from secretome samples by DIA are annotated to extracellular region. Results are displayed as percentage of total IDs annotated to each GO cellular compartment with terms.

Table 3.6: Results from overrepresentation test of secretome DIA identifications against human genome. DIA secretome identifications were tested for overrepresentation against human genome using PANTHER GO-Slim Cellular Component annotation database, and the Fisher's Exact with FDR multiple test correction test. Fold enrichment of statistically significant overrepresented terms, with FDR <0.05 are shown with associated p-value. Data ranked by fold change.

PANTHER GO-Slim Cellular Component	Fold enrichment	p-value
nuclear envelope (GO:0005635)	3.25	5.33E-03
cell junction (GO:0030054)	3.18	2.53E-03
cytosol (GO:0005829)	2.87	5.69E-10
actin cytoskeleton (GO:0015629)	2.81	3.78E-04
extracellular matrix (GO:0031012)	2.34	4.48E-02
ribonucleoprotein complex (GO:0030529)	2.03	1.72E-03
cytoplasm (GO:0005737)	1.75	1.48E-13
mitochondrion (GO:0005739)	1.73	4.22E-02
macromolecular complex (GO:0032991)	1.60	3.38E-06
protein complex (GO:0043234)	1.54	1.81E-04
intracellular (GO:0005622)	1.52	2.29E-13
cell part (GO:0044464)	1.47	5.27E-12
organelle (GO:0043226)	1.37	2.46E-05
nucleus (GO:0005634)	1.37	9.33E-03
plasma membrane (GO:0005886)	0.65	2.30E-02
integral to membrane (GO:0016021)	0.41	9.27E-04

Table 3.7: Proteins associated with GO term extracellular matrix, implicated in overrepresentation of this term. DIA secretome identifications were tested for overrepresentation against human genome using PANTHER GO-Slim Cellular Component annotation database, and the Fisher's Exact with FDR multiple test correction test. Those proteins associated with the overrepresentation of 'extracellular matrix (GO:0031012) are listed here.

Protein Name	Gene Name
Metalloproteinase inhibitor 1	TIMP1
Periostin	POSTN
von Willebrand factor	VWF
Collagen alpha-2(IX) chain	COL9A2
Leucine-rich repeat and WD repeat-containing protein 1	LRWD1
Hemopexin	HPX
Laminin subunit alpha-5	LAMA5
Transforming growth factor-beta-induced protein ig-h3	TGFBI
Collagen alpha-2(I) chain	COL1A2
Laminin subunit beta-1	LAMB1
Neutrophil collagenase	MMP8
Collagen alpha-1(III) chain	COL3A1

Comparing the peak areas of DTT to untreated, etoposide to untreated and DTT to etoposide, t-tests were performed in MarkerView™ to determine the significance score of the fold change of those proteins identified. Only a small fraction of those identified had a p-value score of less than 0.05 to indicate significance (Figure 3.12).

The p-values were adjusted using the p.adjust function in R with the Benjamini-Hochberg method. Without adjustment, a p-value of less than 0.05 indicates that the probability of obtaining the same value being tested (in this case the fold change) by chance, i.e. in a case where the null hypothesis is in fact correct, is 5%. If for each test there is a 5% chance it is a false positive, there may be acceptable numbers of false positives when only a few tests are being performed, but as the number of tests increases, so then does the number of false positives. The Benjamini-Hochberg method for multiple testing describes a way to control the FDR, which enforces a limit on the number of identifications within a set that it is acceptable to be false[167]. This method involves ranking the p-value scores from smallest to largest and comparing these to the Benjamini-Hochberg critical value, defined as $(i/m)Q$ where i is the rank, m is the total number of tests and Q is the FDR chosen. When using this method to adjust the p-values, P is multiplied by m/i, or the adjusted value for the next higher p-value, whichever is smallest. A result is deemed significant if the adjusted p-value is smaller than the desired FDR[168]. In this analysis, the FDR has been set at 5%, or 0.05.

Following p-value adjustment, there were three proteins identified as significant after multiple test correction when comparing DTT treatment to Etoposide, although only one with a fold change greater than 2 (Figure 3.12; Table 3.8). However, these were proteins that were unexpected. G-rich sequence factor 1 (GRSF1; fold change: 6.99; adjusted p-value: 0.03) is involved in the regulation of post translational mitochondrial gene expression and was greatly increased in the DTT treated sample compared to etoposide treated. This is a mitochondria/cytoplasm located protein and, although it is possible that alternatively spliced forms of this protein may be unconventionally secreted, it is more likely that it is present in the media sample collected due to cell death. Bcl2 like protein 1 (BCL2L1; fold change: 0.81; adjusted p-value: 0.047) is an inhibitor of cell death, slightly increased in the etoposide treated sample compared to DTT. This suggests that etoposide may induce more of a protective response than DTT which may be more apoptotic. It would also corroborate with etoposide driving cells to senescence rather than apoptosis. Finally, Neurobeachin-like protein 2 (NBEAL2; fold change: 1.59; adjusted p-value: 0.047) is a ubiquitous ER localised protein that is known to be involved in thrombopoiesis and platelet biogenesis, however it has also been implicated in dermal wound healing and myofibroblast differentiation. Deppermann et al., showed that NBEAL2 knockout mice had a reduced ability to heal dermal wounds due to reduced ECM secretion[169]. NBEAL2 was implicated in the correct formation of mature TGF-β by platelets that is released on wounding to trigger healing events including fibroblast differentiation and ECM deposition[169]. DTT treatment appeared to increase the release of this protein, which may suggest a role for reductive stress in the triggering of wound healing amongst a damaged population of fibroblasts.

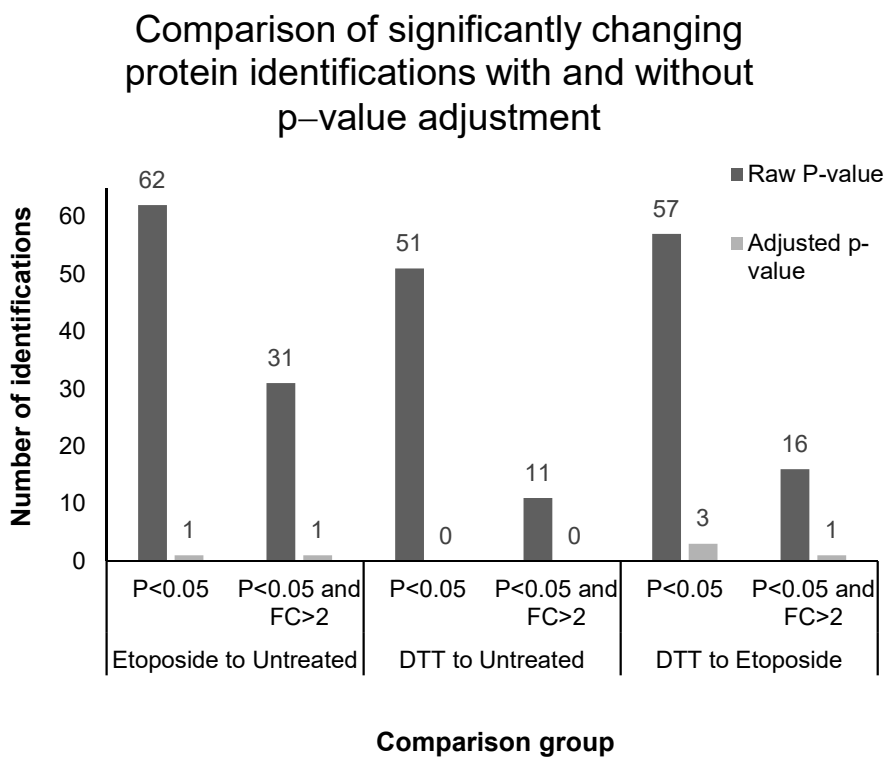


Figure 3.12: Comparison of identification counts in different treatment comparisons. The number of identifications deemed significant (p-value < 0.05) were counted in each of three treatment comparisons: etoposide to untreated, DTT to untreated, and DTT to etoposide. ≤ half of these identifications were also found to have a fold change (FC) of greater than 2. These numbers were again compared following adjustment of the p-value for multiple testing, which greatly reduced the counts in all cases.

Table 3.8: Significant proteins in DTT to Etoposide comparison. Gene names, average summed peak area of all fragment ions, fold change and adjusted p-values are given for proteins that differed between DTT and etoposide treatments, where the differences were determined to be statistically significant (Student's t-test, n=3 technical replicates).

Gene Name	Average peak Area		Fold Change	Adjusted p-value
	DTT	Etoposide		
GRSF1	22651.13292	3237.228423	6.997	0.030
BCL2L1	60509.04458	74254.1249	0.815	0.047
NBEAL2	91810.22526	57733.8525	1.590	0.047

3.6 Discussion and Conclusion

This chapter has described the development of methods to investigate the response of human dermal fibroblasts to ER stress at the proteome level. Methods for cell lysis, protein digestion and sample fractionation have been explored, establishing a platform for the robust proteomic analysis of fibroblasts that are discussed in chapters 5 and 6. To analyse the proteome of a group of cells, a lysate is first generated and acetone precipitated to remove TX100. Samples are then digested with trypsin and fractionated using SCX StageTips™. Immunoprecipitation and STRING-db have also been explored as ways to identify protein-protein interactions with some success. Analysis of conditioned media from fibroblasts proved more challenging with FCS masking the identification of proteins using DDA. The use of higher specification mass spectrometers, which provided the opportunity for DIA, greatly improved the number of identifications possible from secretome samples. However, under these experimental conditions, no known markers of senescence could be detected as significantly increased in the conditioned media of either the etoposide (senescence inducing) or DTT treated cells.

The analysis so far has used an FDR of 5%, or 0.05, which permits 5% of the values to be false positives. This is the standard FDR used in biological analysis but is quite a stringent parameter: if further tests were to be performed to validate the identifications it may well be acceptable for 10 or 15% of the hits to be false positives. Adjusting the FDR would increase the number of identifications deemed to be significant and therefore increase the opportunity for analysis, although this would increase the demand for validation of all results, which is beyond the scope of this thesis.

The investigations in this chapter have mostly relied on technical replicates, and not fully controlled for biological variability. In the proteomic investigations described in chapters 5 and 6, biological replicates have been performed, together with technical replicates. By controlling for both sources of variability it is possible to increase the significance of the results. For example, if a set of three peak areas from technical replicates had a single outlier, the mean peak area would be much more affected than if there was a single outlier in a set of three biological replicate and three technical replicate values (nine values in total).

There is some debate in the proteomics community over the best method for quantitative proteomic analysis, and whether this should be done using a label-free approach as discussed in this thesis or a label-based approach such as 2D differential gel electrophoresis (2D DiGE) or isobaric tags for relative and absolute quantification (iTRAQ). Quantitation is important for the comparison of data from different treatment groups, and 2D DiGE has traditionally been the most popular method for these types of investigation. However, the advent of gel-free approaches provides a clear advantage in the ability to retrieve a complete and unbiased data set that can be

re-interrogated at any time in the future, to answer new questions without generating new samples. Proteins on a 2D DiGE gel that has been stored for months or years may have degraded to some extent and will have been subject to increased contamination from its storage environment. Identification of spots selected on return to a gel may therefore be more difficult than simply re-interrogating a data set stored on a computer.

Labelling approaches are often considered more accurate than label-free techniques such as SWATH, but they require more expensive reagents, and there is a limit to the number of different samples that can be run simultaneously[170]. In addition, the use of labels which may require metabolic or chemical adjustments to proteins within the cells while they are still growing may mean that label-based quantitation methods are not appropriate for all experimental scenarios and labelling efficiency is not uniform across all proteins[170].

A 2009 study by Patel et. al. analysed the cytoplasmic proteome of *Methylocella silvestris* bacterium using three different quantitative approaches to compare the results[171]. The methods used were gel-based, iTRAQ, and label-free DIA. There was significant overlap between all three methods; however, each method also identified its own unique set of proteins that were not seen with any other technique[171]. This suggests that no one method for quantitation of mass spectrometry based proteomic data is perfect, and a balance must be made to adopt the technique that makes best use of the resources available to obtain suitable data for the experiment in hand.

An important consideration is the amount of sample required for analysis. Whilst a gel-based method such as 2D DiGE may require tens or hundreds of micrograms of total protein, SWATH analysis requires a maximum of only 4 µg per technical replicate. Label-free quantitation approaches are reported as being among the least accurate because both systematic and non-systematic errors are reported in the final data[172]. However, these errors can be controlled for by using sufficient biological and technical replicates. In addition, the dynamic range of SWATH data is much greater than with label-based quantitation techniques which may allow for the quantitation of larger changes between experimental conditions[172].

Ultimately, there are arguments for and against all methods of quantitation in proteomic experiments and it is unlikely that there is a perfect solution to all the issues. The SWATH method used for the proteomic investigations discussed in this chapter has proved an invaluable tool in the development of protocols to analyse the secreted proteome with increased sensitivity and lack of bias. In the remaining chapters of this thesis, the data that SWATH analysis provides demonstrates the ability of this label-free quantitation approach to give a useful and obtainable balance between all the factors that may influence the results. It provides the facility to obtain both a broad understanding of the dermal fibroblast proteome, as well as to investigate specific responses to ER and reductive stress.

Chapter 4: Reductive stress induces a novel signalling response in human dermal fibroblasts

4 Reductive stress induces a novel signalling response in human dermal fibroblasts

4.1 Introduction

It is well documented that fibroblast cells enter a quiescent state when exposed to situations of contact-inhibition or confluency: a reversible cell cycle arrest where proliferation and their secretory profile is reduced[173–175]. During the culture of dermal fibroblast cells on 2D plastic surfaces there is likely to be a substantial proportion of cells in a quiescent state once confluency has exceeded 50%. However, due to their rapid growth time and limited levels of protein recovery from sub-confluent cultures, it is usually necessary to grow cells to confluency levels above 50% to allow for experimental procedures and obtain adequate protein yields in cell lysates. *In vitro* it is possible to stimulate fibroblasts to resume proliferation and secretion by exposure to growth factors including epidermal growth factor (EGF), basic fibroblast growth factor (b-FGF) and platelet derived growth factor (PDGF)[176]. In the studies in this Chapter, PDGF was used to stimulate fibroblasts as the cell biology and dose response of the cells to this treatment had been well characterised by P&G.

PDGF was first described in 1974 by Ross et. al., who observed that platelet containing serum, but not platelet poor plasma, was essential for the promotion of proliferation in smooth muscle cells[177]. It is now known that there are four polypeptide chains within the PDGF gene family, the products of which produce four cysteine-knot-type growth factors PDGF A, B, C and D, the last two forms of which have only been discovered in the last 20 years[178] . PDGFs are active as homo- or hetero-dimers of the following combinations of polypeptide chains: AA, BB, AB, CC and DD. All the polypeptides are synthesised as precursors, which must be proteolytically processed prior to dimerization and activation. It is believed that this processing occurs prior to secretion with A and B chains[179], but C and D chains are thought to be processed within the extracellular space[180].

Dimerised PDGF binds to one of the two PDGF receptors (PDGFR): PDGFRα or PDGFRβ, which are class III receptor tyrosine kinases (RTK), bringing about receptor dimerisation such that one PDGF dimer is bound to a PDGFR dimer. Receptor dimerisation may produce homo- or hetero-dimers in the form $\alpha\alpha$ which may bind AA, BB, CC, or AB PDGF; $\beta\beta$ which may bind BB or DD PDGF or $\alpha\beta$ which may bind BB or AB PDGF[178]. Following receptor dimerisation, conformational changes release autoinhibition of the intracellular kinase domain allowing for the autophosphorylation of intracellular tyrosine residues. These phosphorylated tyrosine residues then permit the docking of a range of adapter molecules which form part a complex signalosome to control cell fate[181] (Figure 4.1).

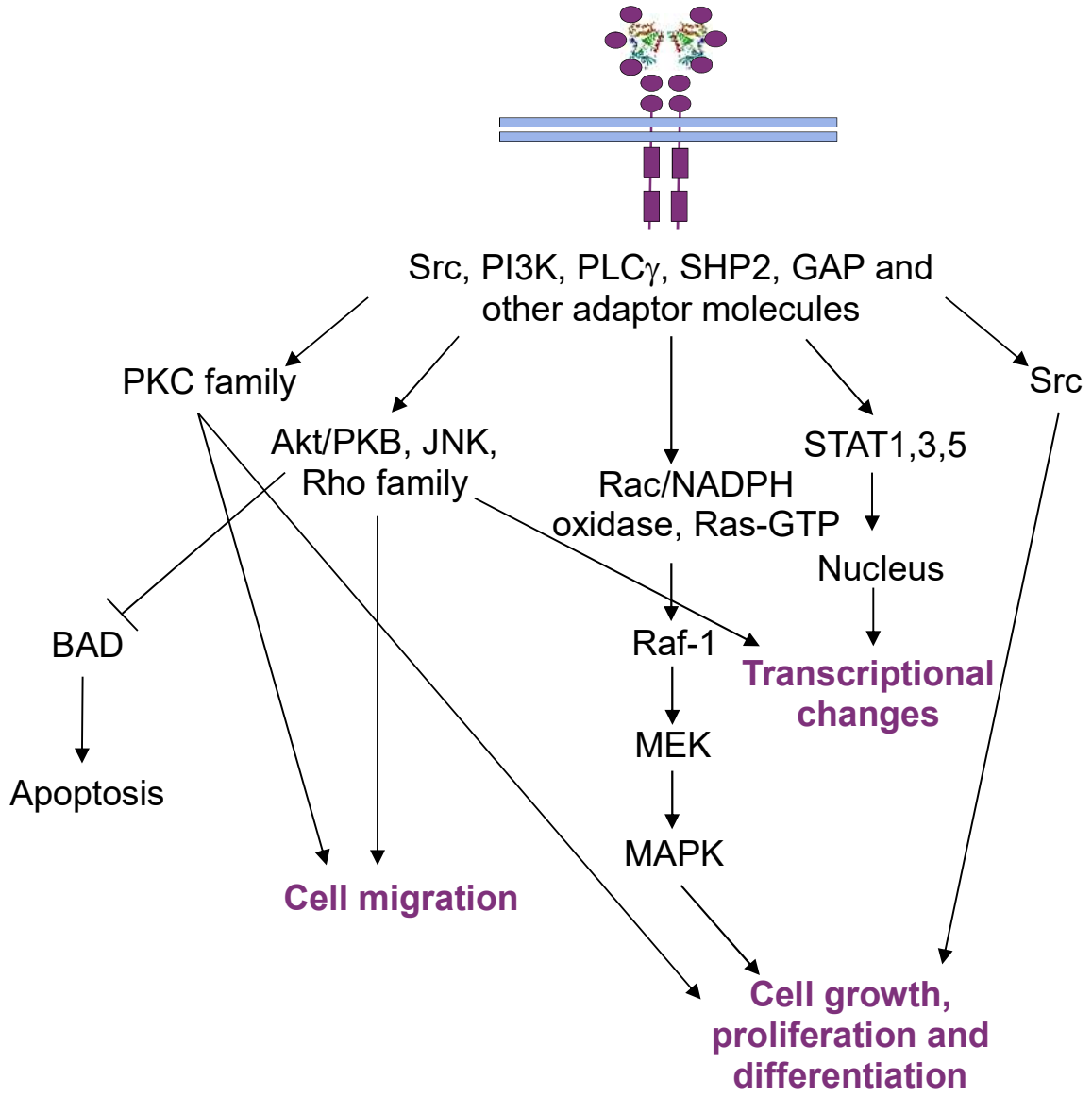


Figure 4.1: PDGF Signalling pathway. Binding of PDGF induces receptor dimerisation and the resulting conformational changes in PDGFR allow autophosphorylation and the binding of intracellular adaptor molecules that trigger downstream signalling cascades. These signalling events ultimately lead to increases in cell migration, growth, proliferation and differentiation amongst a range of transcriptional changes.

PDGF signalling is primarily involved in the promotion of cell growth and proliferation, as well as the inhibition of apoptosis, actin reorganisation and cell migration (Figure 4.1). These signalling events are also important for the correct development of various organs and tissues (PDGFRα) and for angiogenesis and haematopoiesis (PDGFRβ)[182]. It is therefore not surprising that extended or excessive signalling from these molecules is implicated in a range of diseases that involve undue cell proliferation including, for example, cancer and fibrosis[178]. In the past few years there has also been growing evidence for the role of PDGFRs in the autoimmune disease systemic sclerosis (SSc) for which fibrosis is the major symptom[183]. This has led to investigations into the use of PDGFR inhibitors as a treatment for fibrosis. For example, treatment of SSc fibroblasts with a PDGFR inhibitor Imatinib prevents excessive ECM deposition and also inhibits SSc-like fibrosis in mice[184]. It

is therefore also possible that proteomic investigations into the downstream effects of PDGF signalling may provide useful insight into other potential drug targets for the treatment of disease.

Alterations to the redox state of cells have also been implicated in disease states such as SSc and other fibrotic disorders. The importance of redox homeostasis for correct ER function has already been discussed in the introduction to this thesis and much research has been done on the impact of oxidative stress on fibrotic diseases (reviewed by Gabrielli et. al.[185]). However, comparatively little is known about the impact of reductive stress. The highly secretory function of fibroblasts may well make them particularly sensitive to redox imbalances, which could come from conditions of either oxidative or reductive stress. Reductants, which may impose a reductive stress upon the skin, are prevalent in everyday life: sulfur dioxide pollutes our air, and thioglycolate is the major component of depilation creams and hair permanents or straightening treatments. It is therefore of interest to investigate the specific effects of reductive stress on dermal fibroblast cells.

DTT is a potent reducing agent first described by Cleland in 1964 as having a very low redox potential, and able to reduce all disulfides[186]. Reduction of disulfide bonds happens through two sequential reactions, with the concurrent oxidation of DTT into a cyclic compound (Figure 4.2). This chemical can cross membranes and inhibit formation of disulfide bonds within the ER leading to significant stress within cells[187]. The effects of DTT in cell culture have been well characterised, with treatment times of as little as 10 minutes inducing an ER stress response[188].

This chapter will describe the signalling responses of BJ fibroblast cells to PDGF with and without reductive stress induced by DTT. The potential consequences of these responses with respect to cell viability and proliferation will also be discussed to gain insight in to the physiological relevance of these findings.

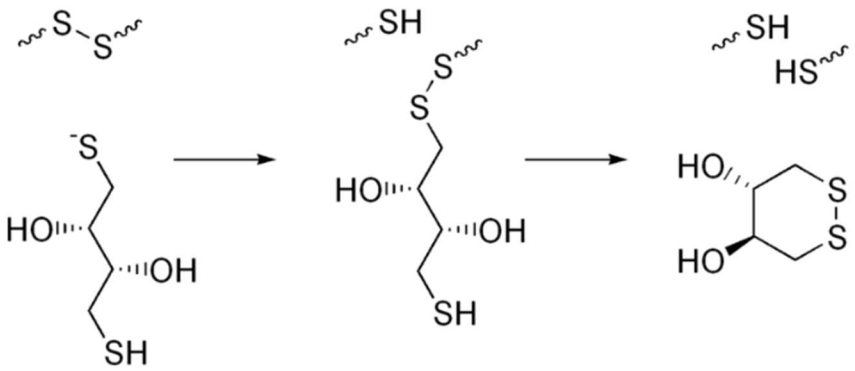


Figure 4.2: Mechanism of action of DTT. DTT reduces disulfide bonds with the concomitant formation of a cyclic disulfide through two reaction steps. All biologically accessible disulfides are reduced by this action.

4.2 Determination of a response to stimulation with PDGF-bb

Before characterising the fibroblast responses to reductive stress, establishing the response of the cells to PDGF stimulation was first required. Due to the presence of growth factors in FCS, cells were treated with PDGF under serum free conditions so that any growth factor-associated responses could be deemed specific to the PDGF treatment. The PDGF B isoform was used (PDGF-bb) to treat the cells so that results were comparable to data held by P&G. Human dermal fibroblasts (HDF) were first used to establish the protocol, and their response to PDGF-bb examined by immunofluorescence and western blot. The protocol was then transferred to the BJ fibroblasts used in the remaining experiments of this thesis, and the expected signalling responses confirmed by western blot.

4.2.1 Response of human dermal fibroblasts to PDGF-bb

Immunofluorescence was used to investigate whether there was any change to the ER morphology following stimulation of fibroblasts with PDGF-bb (Figure 4.3). HDF cells were seeded onto cover slips and treated for 24 hours with media containing 10 % FCS, serum free media or serum free media containing 10 ng ml^{-1} PDGF-bb. Cells were then fixed with paraformaldehyde and permeabilised with TX100 before staining for PDI (red), and the nuclear stain DAPI (blue). PDI was used as this ubiquitous ER resident protein gives an indication of the gross morphology of the ER. No substantial differences were seen in the ER morphology between the three conditions, confirming that the serum free/PDGF treatment was not leading to any unexpected changes to cell organelles.

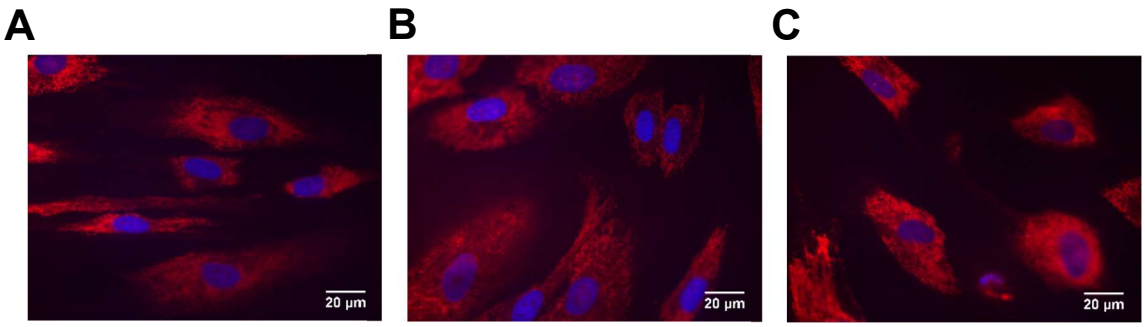


Figure 4.3 Culturing HDFs in serum free media with or without PDGF-bb does not affect the gross morphology of the ER. nHDF cells were grown on cover slips and transferred to serum free media for 24 h with or without PDGF-bb, before fixing and staining for PDI (red) and nuclei (DAPI - blue). Images taken using Zeiss Apotome microscope at 63 x magnification. **A.** Cells grown in media + 10% FCS. **B.** Cells grown in serum free media 24 h. **C.** Cells grown in serum free media 24 h and PDGF-bb.

Western blotting was next used to confirm that the growth factor treatment elicited a response from the cells. HDFs were grown to 70% confluency before treatment with 10 ng ml^{-1} PDGF-bb, in serum free media, for either 10 minutes, or 24 hours after which they were lysed with RIPA buffer containing protease and phosphatase inhibitors (Figure 4.4). RIPA buffer was used because the combination of SDS, sodium deoxycholate, and TX100 allowed for maximum recovery of all protein subpopulations, including those that are nuclear resident. Lysates were analysed by SDS-PAGE followed by western blotting with the PathScan® PDGFR Activity Assay cocktail to detect: Phosphorylated- (p-) forms of PDGFR, SHP2, Akt, p44/42 MAPK (Erk1/2) as well as eIF4E as a loading control.

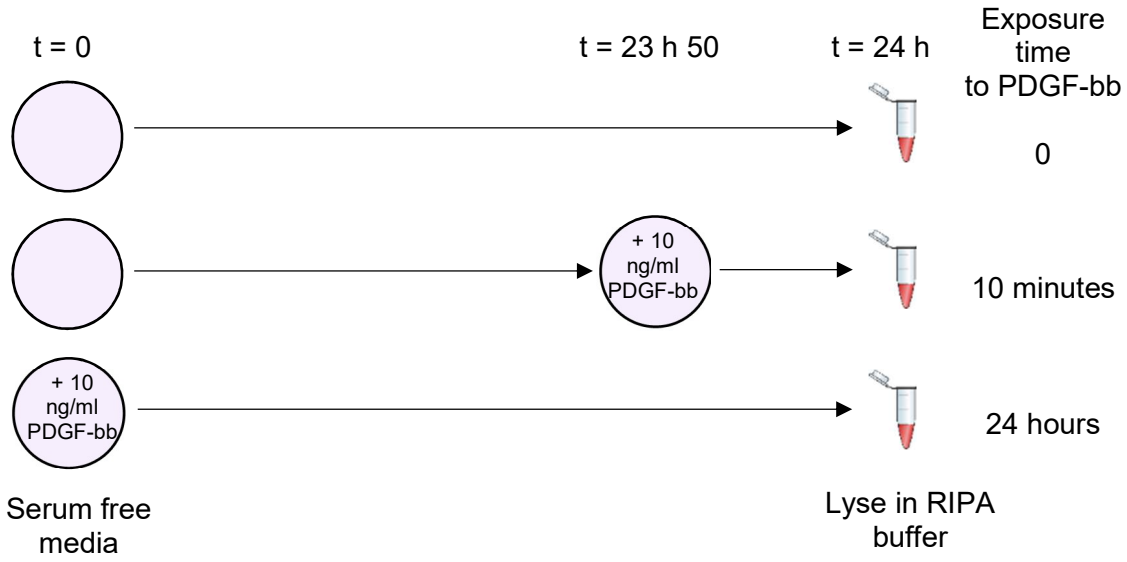
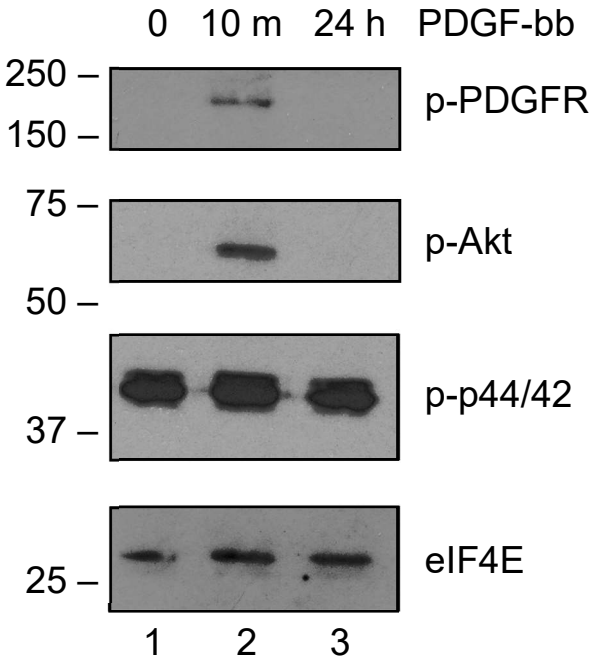


Figure 4.4: Scheme of treatment of cells with PDGF. When fibroblast cells reached approximately 70% confluency they were transferred to serum free media. At this time 10 ng/ml PDGF-bb was also added to the sample to be exposed to PDGF for 24 hours. 10 ng/ml PDGF-bb was added to a further dish 10 minutes before the end of treatment (at t= 23 h, 50 m) and a final dish left without exposure to PDGF. At t=24 h, all cells were lysed in RIPA buffer.

At 10 minutes an obvious response to PDGF-bb was seen, with PDGF receptor phosphorylation at Tyr751, as well as phosphorylation at Ser473 of its downstream effector Akt (Figure 4.5, lane 2). After 24 hours stimulation, the phosphorylation of these proteins had reduced indicating an acclimatisation to the stimulation (Figure 4.5, lane 3). The phosphorylation state of p44/42 remained relatively constant across the three conditions. Thus, an adequate dose of PDGF was used and PDGF signalling responses stimulated within these cells.



.

Figure 4.5 Treatment with PDGF-bb triggers a signalling response in fibroblasts. Lysates from normal human dermal fibroblasts treated with 10 ng ml^{-1} PDGF-bb for 10 minutes or 24 hours in serum free media were analysed by 10% SDS-PAGE followed by immunoblotting with PDGFR activity assay western antibody cocktail containing antibodies against p-PDGFR, p-Akt, p-p44/42 and eIF4e.

4.2.2 The signalling response to PDGF stimulation is similar in BJ fibroblasts

The same treatment set up described in section 4.2.1 (Figure 4.4) was then used to test the signalling responses to PDGF-bb in BJ fibroblasts. A very similar response was seen, with phosphorylation of the PDGF receptor and Akt appearing at 10 minutes (Figure 4.6, lane 2). However, at 24 hours there was still a signal from p-Akt, although less that at 10 minutes, which was not seen in the HDF treatments (Figure 4.6, lane 3). This suggests that the BJ fibroblasts may acclimatise to stimulation less quickly than HDF cells and therefore, whilst responses are likely to be broadly similar there will likely be some variability in responses between different cell lines. For simplicity, and comparability with data from P&G, BJ fibroblasts were used for the remaining experiments in this thesis.

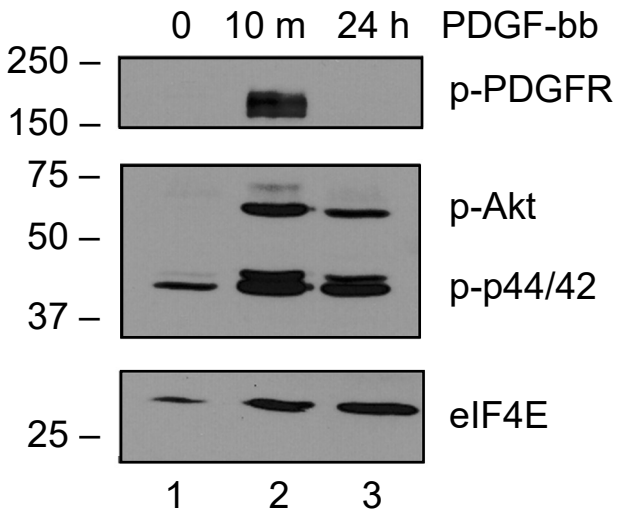


Figure 4.6 Treatment with PDGF-bb also triggers a signalling response in BJ fibroblasts. Lysates from BJ fibroblasts treated with 10 ng ml^{-1} PDGF-bb for 10 minutes or 24 hours in serum free media were analysed by 10% SDS-PAGE followed by immunoblotting with PDGFR activity assay western antibody cocktail containing antibodies against p-PDGFR, p-Akt, p-p44/42 and eIF4e. Representative image of 3 biological replicates.

4.3 Effects of reductive stress on PDGF signalling pathways

Having established that PDGF signalling was effective in these cells, it was then possible to investigate the additional effects of reductive stress on PDGFR signalling pathways. DTT treatment was given for 1 or 6 hours to BJ fibroblast cells at approximately 70 % confluency, on a background of 24 hour serum free treatment with, or without PDGF. Following treatment, cells were washed in PBS and lysed with RIPA buffer (Figure 4.7). This provided a way to compare how DTT affected signalling responses in quiescent cells (SFM only) versus cells activated with PDGF treatment.

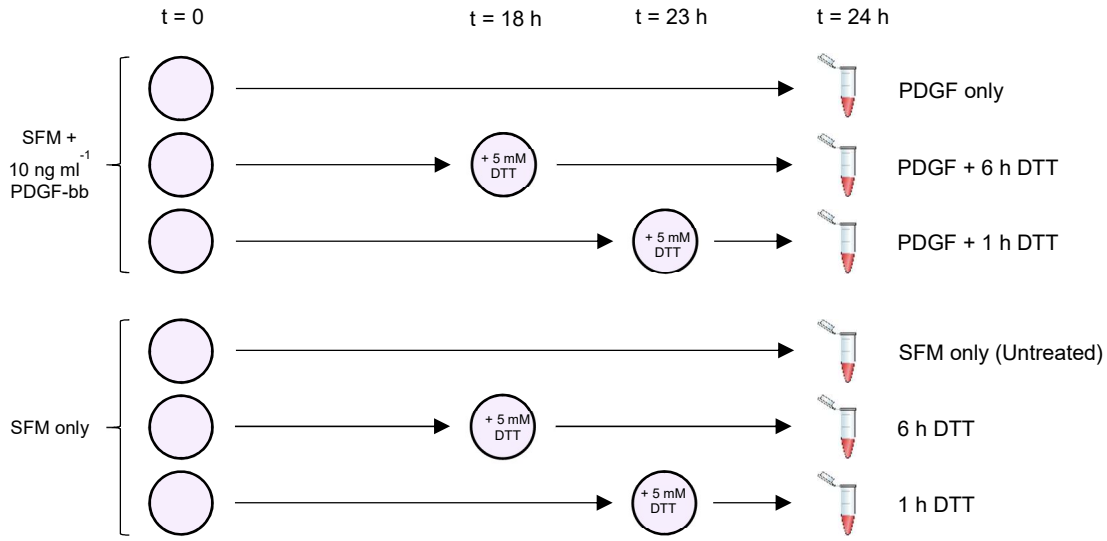


Figure 4.7: Scheme of treatment of cells with DTT ± PDGF. When BJ fibroblasts reached approximately 70 % confluency they were transferred to serum free media with or without 10 ng ml^{-1} PDGF. After 24 hours under these conditions cells were washed and lysed in RIPA buffer. For DTT treatments, 5 mM DTT was added to cells either 1 or 6 hours prior to lysis.

4.3.1 DTT induces an ER stress response in BJ fibroblasts

DTT is well known to induce an ER stress response in many cell types. To demonstrate that this was also the case in the BJ fibroblasts, two assays were performed for markers of ER stress: reverse transcription-polymerase chain reaction (RT-PCR) and *PstI* digestion of XBP1 cDNA; and an immunoblot of cell lysates with an antibody against phosphorylated-eIF2α. The effects of DTT treatment on ER stress markers was assessed both in the presence and absence of PDGF to evaluate whether PDGF modulated the stress response.

XBP1 is a transcription factor that, once activated, binds to DNA and initiates the transcription of UPR genes. Activation occurs through the splicing and removal of a hairpin loop region of the XBP1 mRNA by IRE1 (Figure 1.7), which contains within it a specific digestion site for *PstI*[189,190]. Thus, under normal conditions it is possible to visualise two bands on an agarose gel of XBP1 RT-PCR products following *PstI* digestion, whereas under stressed conditions, the *PstI* digestion site has been removed and only one band is visible (Figure 4.8A). To observe XBP1 splicing in BJ fibroblasts,

cells were treated as described above (section 4.3) and lysed in TRI reagent. RNA was extracted from TRI reagent lysates using chloroform precipitation and subject to RT-PCR with forward and reverse primers for XBP1. RT-PCR using the same RNA extracts was also performed with primers for β-actin to act as a loading control. Prior to visualisation of products on a 2 % agarose gel, XBP1 RT-PCR products were incubated for 2 hours with *PstI* at 37 °C. As expected, a single band was visible in all samples treated with DTT consistent with the concept that DTT induces ER stress (Figure 4.8B,C).

Phosphorylation of eIF2α is also known to increase under conditions of ER stress (Figure 1.8). eIF2α is a subunit of the eukaryotic translation initiation factor 2, which is required for the formation of the 43S translation initiation complex made up of small ribosomal subunit, Met-tRNA, GTP and eIF2. When eIF2α is phosphorylated this 43S translation initiation complex is unable to assemble and protein synthesis is inhibited[191]. To test the phosphorylation state of eIF2α, cell lysates were prepared from treated BJ fibroblasts using RIPA buffer with protease and phosphatase inhibitors, and these were subject to separation by SDS-PAGE, transfer to PVDF membrane and immunoblotted for p-eIF2α and β-actin as a loading control. The signal for phosphorylated eIF2α was increased in DTT treated samples providing further evidence for the induction of ER stress (Figure 4.8D,E). Neither eIF2α phosphorylation or XBP1 splicing was altered by co-treatment with PDGF (Figure 4.8B,D).

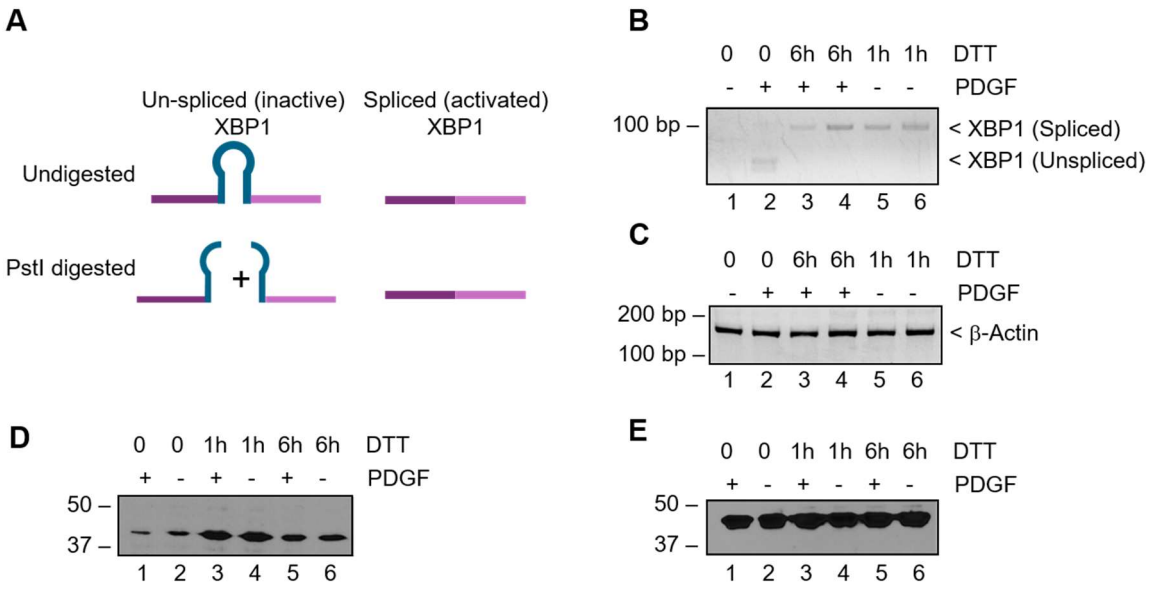


Figure 4.8: DTT induces ER stress in BJ Fibroblasts. A. Schematic describing the expected PCR products of XBP1 depending on activation status. *PstI* restriction endonuclease site exists only in the hairpin loop of the un-spliced (inactive) form of XBP1. When XBP1 is activated, this restriction site is removed and *PstI* digestion no longer occurs. **B.** RT-PCR products from XBP1 assay and **C.** β-Actin was visualised with ethidium bromide in 2 % agarose gel. **D.** Immunoblot for phosphorylated eIF2α and **E.** β-Actin.

4.3.2 DTT induces growth factor-independent phosphorylation of Akt

As expected from the initial assessment of signalling following PDGF activity, the lysates from cells treated for 24 hours in serum free media ± PDGF only (i.e. no DTT treatment) displayed phosphorylation of p44/42 (ERK1/2) but only very weak phosphorylation of Akt (Figure 4.9A,B, lanes 1 and 2). The same signalling pattern was also seen after 1 hour DTT in both + and – PDGF cases (Figure 4.9A,B, lanes 3 and 4). Surprisingly, however, after 6 hours DTT there was a strong signal for p-Akt and the signal for p-p44/42 appeared greatly diminished. The same signalling pattern was seen both in the presence and absence of PDGF stimulation (Figure 4.9A,B, lanes 5 and 6). This demonstrates a signalling pattern that has not been previously described whereby DTT treatment stimulated growth factor-independent activation of the Akt pathway, with concomitant de-activation of the p44/42 pathway.

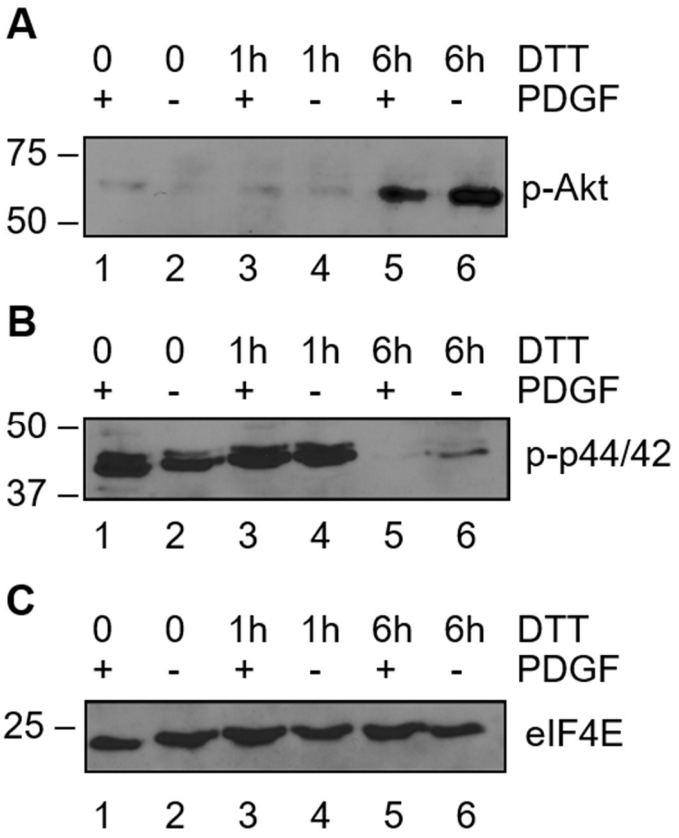


Figure 4.9: DTT stimulates Akt phosphorylation with or without PDGF. Lysates from BJ fibroblasts treated with 0, 1, or 6 hours 5 mM DTT in serum free media with or without 10 ng ml^{-1} PDGF-bb were analysed by 10 % SDS PAGE and immunoblotted with a western blotting antibody cocktail containing antibodies against **A.** p-Akt **B.** p-p44/42 **C.** eIF4e. Representative blot of n=5 biological replicates.

The Akt pathway is well known to be involved in many pro-survival mechanisms including cell growth, proliferation, inhibition of apoptosis, angiogenesis, metabolism and even migration and invasion (reviewed by Manning and Cantley in 2009[192]). p44/42 are members of the MAPK family and participate in the Ras-Raf-MEK-ERK signalling cascade. The targets of this signalling pathway are wide and in some cases opposing, including cell migration, cell cycle progression and differentiation but also cell death and apoptosis under certain conditions[193]. The contrasting phosphorylation states of these two signalling pathways suggests that a persistent reducing environment requires the cell to carefully balance the pro-survival stress response with a need to remove those cells which are beyond repair.

4.3.3 Signalling response is unique to redox stress

To test whether the Akt/p-44/42 response is a result of ER stress in general, or specifically results from redox stress, cells were treated with thapsigargin. Thapsigargin blocks the ATP-dependant calcium pump in the ER, leading to loss of calcium homeostasis. This dysregulation of the usually tightly controlled calcium levels triggers ER stress and the UPR in a way that is not redox related[194].

Cells were treated with 2 μM thapsigargin for 1 or 6 hours in serum free media with or without PDGF before being lysed in RIPA buffer, and lysates were immunoblotted for p-Akt, p-p44/42 and β-actin as a loading control. Whereas with DTT treatment a substantial increase in Akt phosphorylation was seen at 6 hours, with thapsigargin there appeared to be an increase in phosphorylation levels with the 1 hour treatment, which then returned to the initial levels at the 6 hour time point. This suggests that the Akt phosphorylation occurs earlier in the cellular response to thapsigargin than in the cellular response to DTT. In contrast to the results seen with DTT, the signal from p-p44/42 increased with thapsigargin treatment where it had previously decreased with DTT (Figure 4.10). Together, these results demonstrate a uniqueness to the phosphorylation events seen following DTT treatment suggesting a redox-specific response within the cells.

4.3.4 Other reductants illicit similar, but not identical, signalling responses

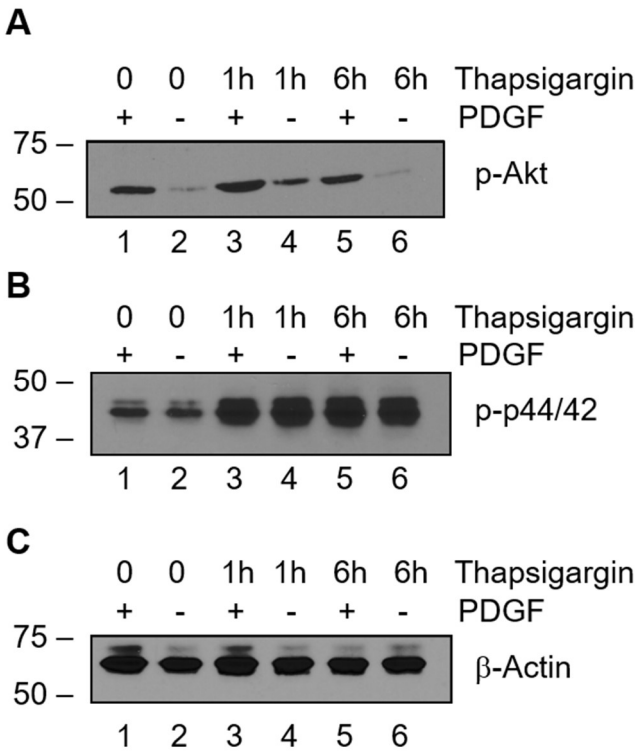


Figure 4.10: Thapsigargin does not stimulate the same signalling responses as DTT. In contrast to the responses seen with DTT treatment, thapsigargin treatment elicits **A.** phosphorylation of Akt that peaks at 1 hour before reducing again by 6 hours, and **B.** increased p44/42 phosphorylation. **C.** β-Actin loading control.

Other chemicals that may affect the redox condition of cells were used to establish whether the phosphorylation of Akt and de-phosphorylation of p44/42 seen was specific to DTT or occurred more generally in response to redox imbalance.

The β-amino acid taurine, also known as 2-aminoethanesulfonic acid, is a potent anti-oxidant that occurs naturally from methionine and cysteine metabolism. It is known to be involved in many functions within the body including anti-oxidation and detoxification[195] and has also been shown to have anti-fibrotic properties[196–198]. Whilst taurine is found naturally within the body, these anti-fibrotic characteristics have led to its use in cosmetics such as skin treatments designed to reduce scarring and anti-ageing creams. Thioglycolate is present in many depilatory creams and lotions used for cosmetic hair removal[199] and works by reducing disulfide bonds in keratin such that the hairs are weakened and easily removed. Thioglycolate salts are also commonly used in hair permanent and straightening treatments. Finally, hydrogen peroxide is a well-known chemical that induces oxidative stress in cells and has previously been shown to induce a peak of Akt phosphorylation with short term treatment that then reduces to basal levels after long-term exposure[200]. This induces an oxidative, rather than reductive, stress in cells and was used to compare the signalling responses in this cell type to determine whether they were part of a redox or specifically reductive stress response.

BJ fibroblasts were grown to 70% confluency before being treated with 5 mM taurine, 5 mM potassium thioglycolate, or 5 mM hydrogen peroxide for 10 minutes, 1 or 6 hours. Cells were then lysed in RIPA buffer and the lysates analysed by immunoblot for p-Akt and p-44/42. As expected from the literature, hydrogen peroxide treatment induced a rapid and significant increase in Akt phosphorylation within 10 minutes and this declined following long-term exposure (Figure 4.11A)[201]. The same pattern of phosphorylation was seen with p-44/42 (Figure 4.11B). With taurine treatment, a p-Akt signal was seen at all time points (Figure 4.11C) and p-44/42 phosphorylation reduced after an initial increase at 10 minutes (Figure 4.11D). Though these responses are not as well-defined as seen with DTT treatment, they do demonstrate a similar pattern. Thioglycolate induced a pattern of signalling more like hydrogen peroxide than DTT. An increase in phosphorylated-Akt signal was seen at 10 minutes, but this was then greatly reduced after extended exposure (Figure 4.11E) and the same pattern was seen with phosphorylation of p44/42 (Figure 4.11F).

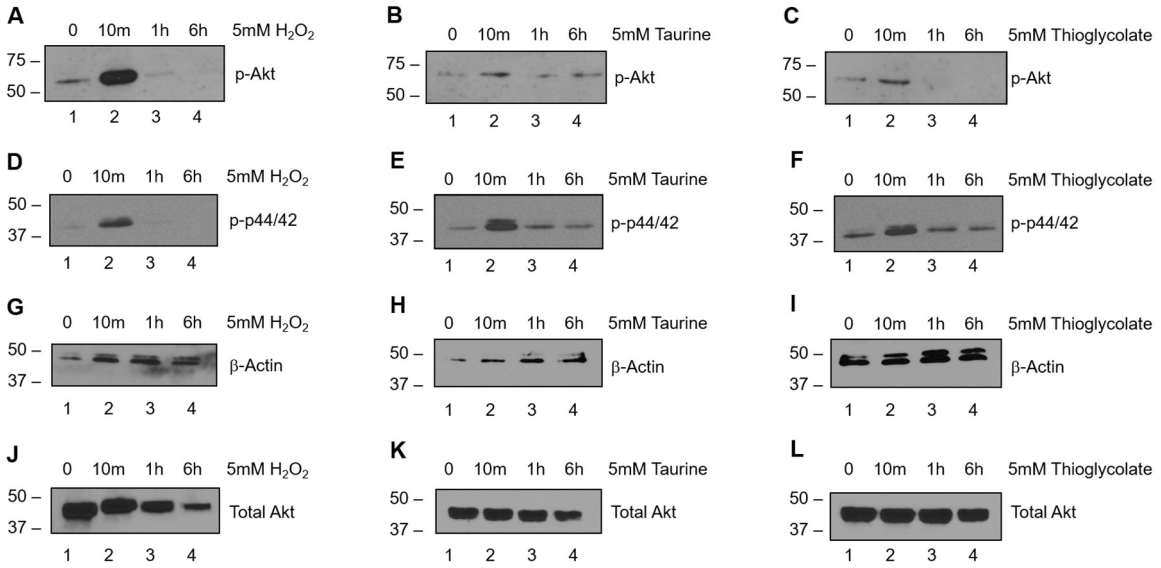


Figure 4.11: Signalling responses to treatments that alter the redox balance of cells. Lysates from BJ fibroblasts treated with 5 mM hydrogen peroxide (**A, D, G, J**), Taurine (**B, E, H, K**) or Thioglycolate (**C, F, I, L**) were analysed by 10% SDS-PAGE followed by immunoblotting for p-Akt (**A, B, C**), p-p44/42 (**D, E, F**), β-Actin (**G, H, I**) or total Akt (**J, K, L**). Representative images from n=2 biological replicates.

These results do not clearly define a specificity of the response to DTT treatment as being reducing or redox specific, and it is possible that the signalling responses seen are the result of a fine balance in the redox state of the cells. The changes to Akt phosphorylation seen with DTT are clearly different to those observed with hydrogen peroxide, suggesting a reductive stress specificity of the response. However, thioglycolate, which is also a reducing agent, does not appear to elicit the same pattern of Akt phosphorylation. Perhaps thioglycolate is less potent than DTT and therefore the redox balance within the cell is disrupted to a lesser extent reducing the requirement for activation of the protective, anti-apoptotic effects of the Akt signalling pathway. However, this would require further testing of a range of concentrations and timepoints. The mode of action of taurine is less well known and is thought to function more generally as an anti-oxidant than specifically as a reducing agent. Jong et. al., described the role of taurine in the mitochondria to protect against superoxide generation by increasing the production of proteins involved in the electron transport chain[202]. However, studies have also implicated taurine in the modulation of intracellular levels of the reducing equivalents NADH and NADPH[203]. This suggests that taurine may affect the overall redox state of the cells by modulating the intracellular production of ROS, as well as the levels of NADH and NADPH. It is possible that the signalling responses seen in response to DTT are not just due to a reductive stress but more a reflection on the level of disruption to the cells redox homeostasis, and the potential ability for the cell to survive this stress.

To further characterise the signalling responses seen following DTT treatment, an extended time course was used to determine the time periods over which Akt phosphorylation was seen.

4.3.5 Extended time course of DTT treatment reveals chronic Akt phosphorylation

BJ fibroblasts were treated for 10 minutes, 30 minutes, 1, 2, 3, 4, 5, or 6 hours with 5mM DTT in serum free media with or without PDGF before blotting for p-Akt and p-p44/42 with total Akt as a loading control (Figure 4.12). Although obviously stronger in the early timepoints, the signal for phosphorylated Akt was present at all time points in the presence of PDGF with the signal weakening from about 3 hours DTT (Figure 4.12A). Conversely, phosphorylated Akt in the absence of PDGF only appeared very weakly at 10 minutes, with the signal becoming gradually stronger between 1 and 3 hours from which point it remained strong (Figure 4.12B). The p-p44/42 signal behaved similarly both with and without PDGF: there was an increase at the 10 minute timepoint, but this was greatly reduced, if not missing completely by 6 hours DTT. With PDGF, the signal began to diminish from 2 hours and by 4 hours was no longer visible (Figure 4.12C) whereas in the absence of PDGF, the signal lasted longer, diminishing only from 3 hours and was only completely absent at the 6-hour timepoint (Figure 4.12D).

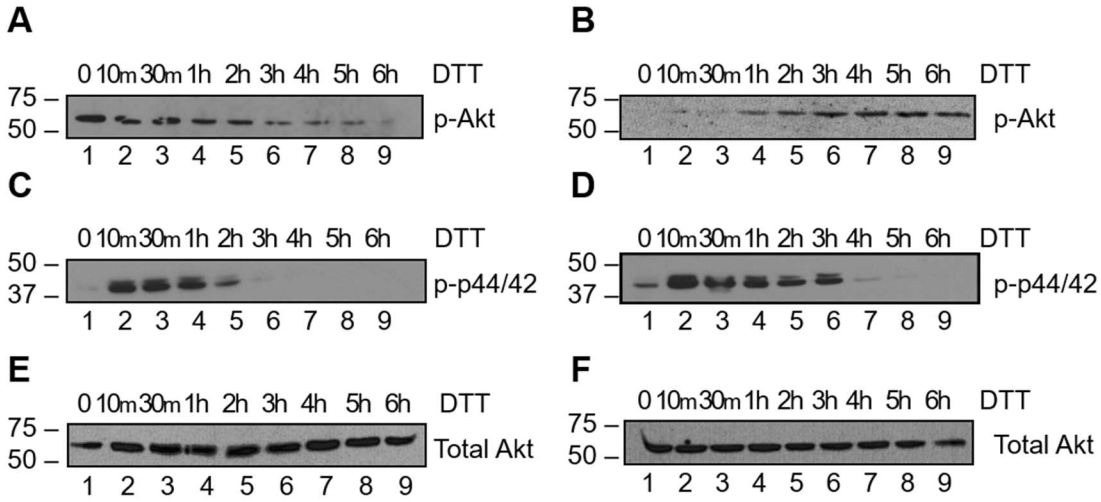


Figure 4.12: Chronic Akt phosphorylation in response to DTT treatment. Lysates from BJ fibroblasts treated with DTT in the presence (**A, C, E**) or absence (**B, D, F**) of 10 ng ml^{-1} PDGF-bb were analysed by 10% SDS-PAGE followed by immunoblotting for p-Akt (**A, B**), p-p44/42 (**C, D**), or total-Akt (**E, F**).

The precise interactions of the signalling responses to PDGF and DTT are not known. It is possible to suggest, however, that after 2 hours of DTT exposure the anti-apoptotic effects of the Akt signalling pathway are no longer sufficient to protect cell survival and are therefore alleviated in PDGF stimulated cells under chronic DTT stress. In addition, it is possible that the reduction in p44/42 phosphorylation is indicative of a reduction in the promotion of cell growth and proliferation, which has an earlier onset in the more sensitive, PDGF treated cells. It is likely that the fate of the cells is determined by a careful balance between these two signalling events and it is therefore of interest to further investigate the response in terms of cell viability and proliferation potential.

4.4 Consequences of chronic Akt phosphorylation

4.4.1 Akt inhibition by perifosine

To test whether the Akt signalling seen in Figure 4.12 was beneficial or detrimental to the cells, an inhibitor of Akt was used and the effect on cell proliferation observed (Figure 4.13). Perifosine is a synthetic alkyl-phospholipid designed to inhibit the erroneous activation of Akt often seen in cancer cells. Studies have shown that perifosine directly blocks Akt activity, preventing both constitutive and induced phosphorylation without affecting upstream components of the signalling pathway[204,205].

BJ fibroblast cells were treated for 0, 1 or 6 hours with 5 mM DTT in serum free media with or without 10 ng ml^{-1} PDGF-bb. Cells were given 5 µM perifosine, 5 minutes before treatments began. Following treatment, cells were lysed in RIPA buffer and the lysates analysed by immunoblot for p-Akt, and p-p44/42 alongside a positive control lysate from a previous experiment (Figure 4.13A, B). Cell proliferation was tested using the Orangu™ assay developed by Cell Guidance Systems. This assay is based on the reduction of the tetrazolium salt WST-8 by viable cells to an orange-coloured formazan dye, where the amount of dye formed is directly proportional to the number of viable cells. WST-8 is added to cells and incubated for 2 hours before reading the absorbance at 450 nm to determine the amount of formazan dye generated. In healthy and proliferating cells, it is expected that the absorbance would increase between 3 and 4-fold over a period of 6 days. BJ fibroblast cells were seeded into 96 well plates and allowed to settle overnight before treatment for 6 hours with 5 mM DTT either alone or with 5 µM perifosine, 5 µM perifosine alone, or 12.5 µM etoposide. Etoposide was used as a positive control as it is known to reduce cell proliferation (see chapter 3). Orangu ™ assay WST-8 reagent was added and incubated for 2 hours prior to reading absorbance at 450 nm on day 0, day 2 and day 4 (Figure 4.13C).

Perifosine treatment successfully abolished Akt phosphorylation across all timepoints with no effect seen on p44/42 phosphorylation (Figure 4.13A,B). The cell proliferation assay demonstrated a small but insignificant increase in cell proliferation of cells treated with DTT and perifosine compared to those treated with DTT alone (Figure 4.13C). Treatment with perifosine alone also seemed to increase cell proliferation compared to the untreated control suggesting that in some cases Akt can indeed impede cell proliferation rather than wholly promoting it. These results suggest that perhaps the chronic Akt phosphorylation seen in these studies is detrimental rather than beneficial to the cells.

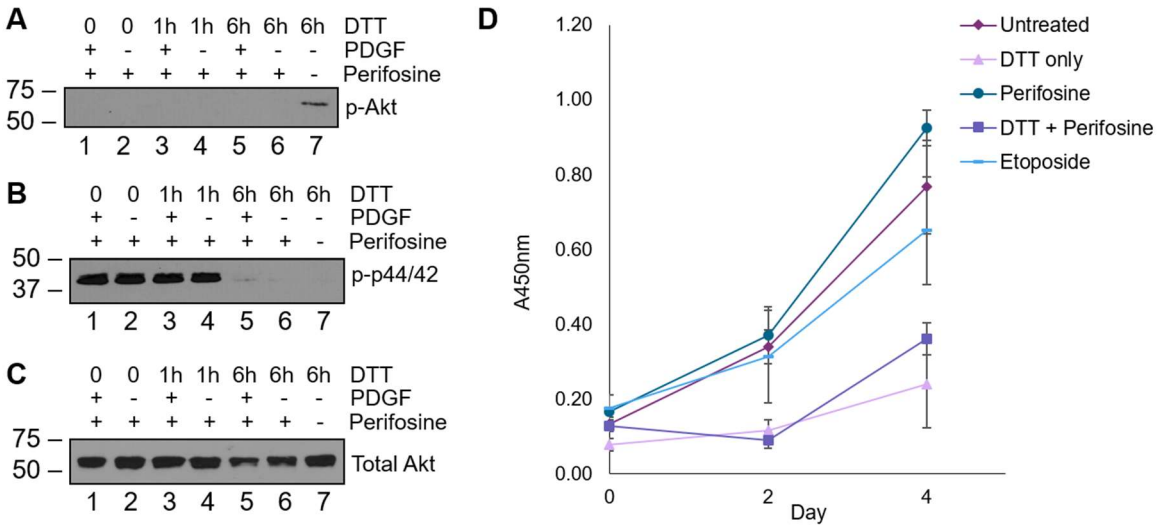


Figure 4.13: Perifosine abolishes Akt phosphorylation. BJ fibroblast cells were given 5 µM perifosine 5 minutes prior to treatment for 0, 1, or 6 hours with 5 mM DTT in serum free media with or without 10 ng ml^{-1} PDGF-bb. Lysates were analysed by 10 % SDS-PAGE and immunoblotted for (**A**) p-Akt, (**B**) p-p44/42 and (**C**) total Akt alongside a positive control from a previous experiment (lane 7). **D.** The Orangu ™ colorimetric assay was used to estimate cell proliferation in response to treatments by incubating WST-8 with cells for 2 hours before reading absorbance at 450 nm at day 0, day 2 and day 4 (n=3).

The phosphorylation of Akt in cells treated with etoposide, which reduced cell proliferation between day 2 and day 4 (Figure 4.13C), was next tested to investigate whether the Akt response was like that seen with DTT. BJ fibroblasts were treated with 12.5 µM etoposide for 10 minutes, 6 or 24 hours and lysed in RIPA buffer. Lysates were immunoblotted for p-Akt, p-p44/42 and total Akt as a loading control (Figure 4.14A, B, C). Long-term etoposide treatment of cells also led to increased Akt phosphorylation and reduced p44/42 phosphorylation similarly to that seen in DTT treated cells.

The proliferation of DTT and etoposide treated cells was also compared. BJ fibroblasts were plated in a 96-well plate and allowed to settle overnight. Treatments of 5 mM DTT or 12.5 µM Etoposide were added for 6 hours before washing and replacing with fresh media. Orangu™ reagent containing WST-8 was added to the cells immediately after treatment was removed (day 0), on day 2 or day 6 and absorbance at 450 nm measured after 2 hours incubation. With both DTT and etoposide treatment, the proliferation of fibroblasts was significantly reduced (Figure 4.14D) and the change in absorbance, which is proportional to the change in number of cells, was also significantly different to the untreated cells (one-way ANOVA ($F(2,6)=30$, $p=0.0008$)) (Figure 4.14E).

Several recent studies have implicated signalling through the Akt pathway, particularly chronic Akt activation, in the development of cellular senescence[206–208]. Senescent cells have a distinct phenotype which was reviewed by Höhn et al. in 2016[209] and described as an increased cell size and protein content, increased number and size of lysosomes and stable growth arrest, as well as a SASP[163]. The SASP is broadly divided into three categories namely, signalling factors such as

interleukins, proteases, and insoluble proteins and extracellular matrix components. Together these factors influence the surrounding tissue microenvironment which often leads to inflammation and in some cases may enable cancer development[163].

Etoposide induces senescence, as discussed in chapter 3, and it was therefore thought possible that chronic reductive stress, as generated through long-term DTT treatment, might also drive cells towards senescence. Senescent cells do not proliferate so if a large proportion of DTT treated cells had undergone senescence, the proliferation rate of the population would be significantly reduced, consistent with the data in Figure 4.14. However, a reduced level of proliferation could be caused by other mechanisms, thus the potential induction of senescence by DTT was tested further.

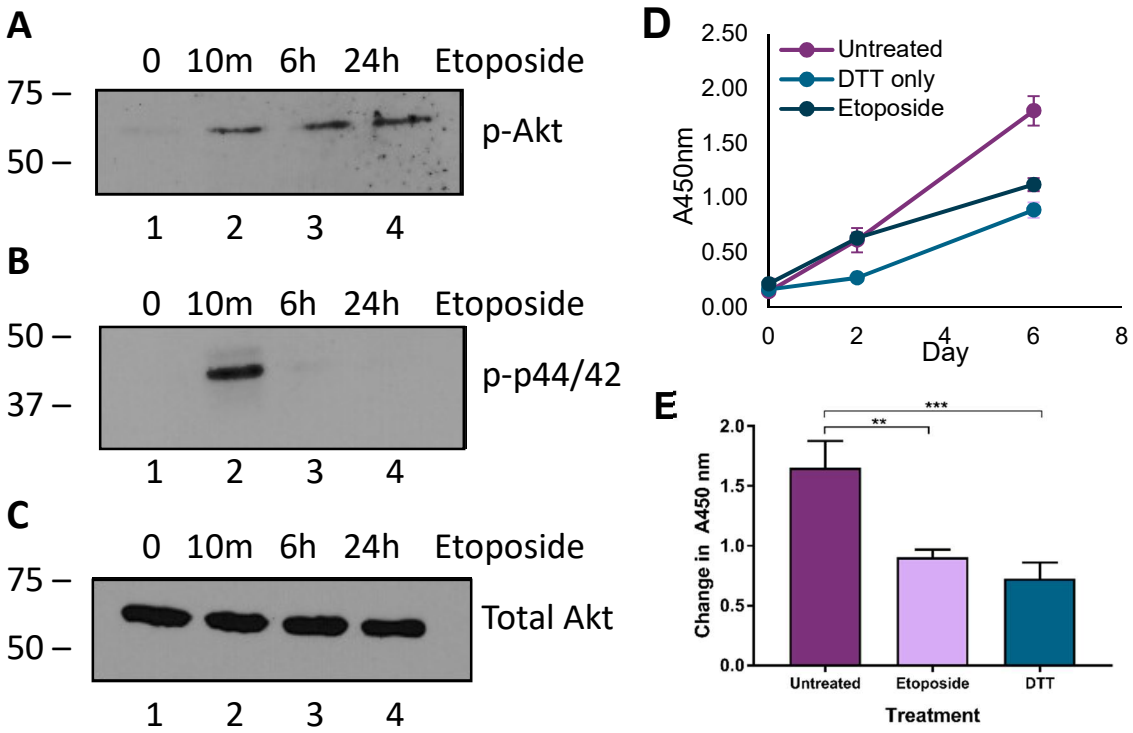


Figure 4.14: Etoposide also induces chronic Akt phosphorylation and both DTT and Etoposide significantly reduce cell proliferation. Lysates from BJ fibroblasts were treated with 12.5 µM etoposide for 10 minutes, 6 or 24 hours were analysed by 10 % SDS PAGE followed by immunoblot for **A.** p-Akt **B.** p-p44/42 and **C.** total Akt. BJ fibroblast cells were also treated for 6 hours in a 96 well plate with either 5 mM DTT or 12.5 µM etoposide. **D.** The Orangu ™ colorimetric assay was used to estimate cell proliferation by incubating WST-8 with cells for 2 hours before reading absorbance at 450 nm at day 0, day 2 and day 6. **E.** The change in absorbance was calculated and a significant difference between means detected using a one-way ANOVA ($F(2,6)=30$, $p=0.0008$). A Dunnett's multiple comparison post-hoc test compared each treatment to untreated and a significant difference was seen with both comparisons. Untreated to Etoposide $p=0.0019$ and Untreated to DTT $p=0.0006$. ** indicates $p<0.01$; *** indicates $p<0.001$.

4.4.2 Senescence

To test for senescence in response to chronic DTT-induced reductive stress a senescence-associated β-galactosidase (SAβGal) assay was used (Figure 4.15). β-galactosidase is a lysosomal enzyme with maximal activity at pH 4 - 4.5. Under conditions of senescence, the cells lysosomal content increases and β-galactosidase activity is detectable at the sub-optimal pH of 6[210]. At pH 6, senescent cells may be stained using the chromogenic substrate X-gal that is converted to a blue precipitate by β-galactosidase.

BJ fibroblast cells were grown to 50% confluency before treatment for 6 hours with 5 mM DTT or 12.5 µM etoposide on a background of 24 hours serum free media. After this time the cells were washed twice with PBS before being returned in to fresh, serum-containing media to recover for 4 days. The cells were then washed and stained overnight with SAβGal staining solution after which they were washed and stored in glycerol before imaging with a light microscope. Following imaging it appeared that, as expected, there were many more blue-stained cells in the etoposide treated sample (Figure 4.15D, E, F) than untreated (Figure 4.15G, H, I). Interestingly, the DTT treated sample (Figure 4.15A, B, C) exhibited blue-staining more like the etoposide sample than the untreated sample suggesting that long-term DTT treatment might indeed induce senescence. However, the results from repetitions of this assay were variable and the background staining from untreated cells could be high, making the results ambiguous. It was also noted that there was no reliable method of quantitation in the SAβGal assay and so determination of differences between treatments relied solely on visual judgements.

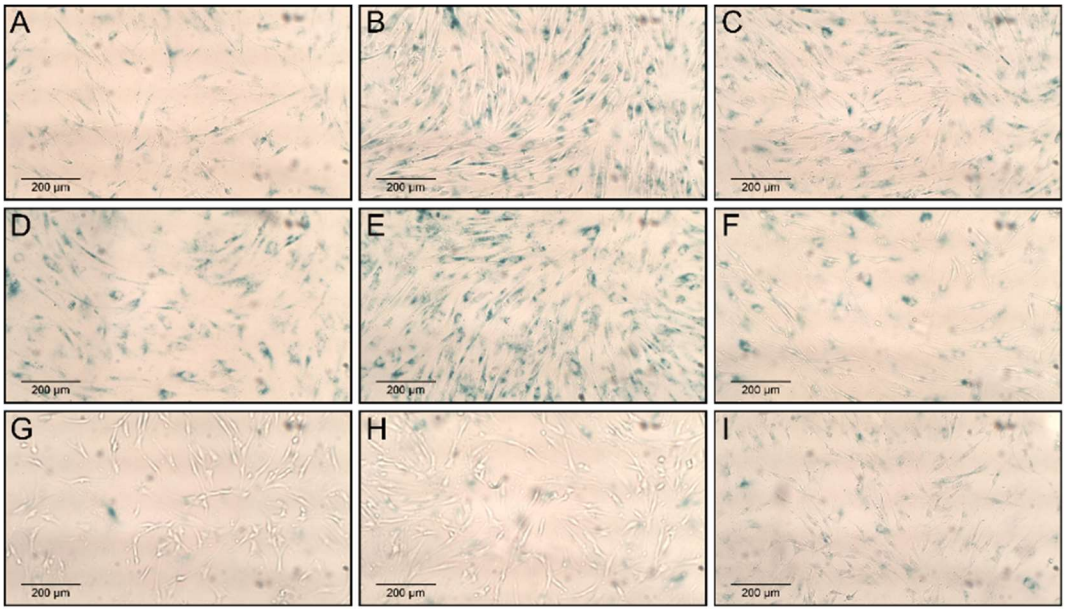


Figure 4.15: Staining for senescence-associated β-galactosidase initially suggests chronic DTT treatment drives cells to senescence. BJ fibroblasts were treated for 6 hours with 5 mM DTT or 12.5 µM etoposide before recovery for 4 days and staining for SAβGal. Three representative images were taken from one biological replicate of DTT treated cells (**A, B, C**), etoposide treated cells (**D, E, F**) and untreated cells (**G, H, I**).

Therefore, this assay was re-visited with the addition of DAPI as a co-stain such that the nuclei of the cells were visible and quantifiable. The samples were then imaged using an inverted wide-field fluorescence microscope with phase contrast filter (Figure 4.16). An additional treatment condition of 5 mM DTT and 5 μM perifosine was also used to help determine whether any effects seen could be reversed by inhibition of Akt phosphorylation.

BJ fibroblast cells on cover slips at 50% confluency were treated with 5 mM DTT ± 5 μM perifosine or 12.5 μM etoposide for 6 hours before being washed and replaced into fresh media to recover. On day 4 after treatment, the cells were washed, fixed and stained for SAβGal as described previously. Following SAβGal staining, the cover slips were washed in PBS and stained with DAPI before mounting on microscope slides. Representative images were then taken from 3 biological replicates of each treatment condition (Figure 4.16A-D).

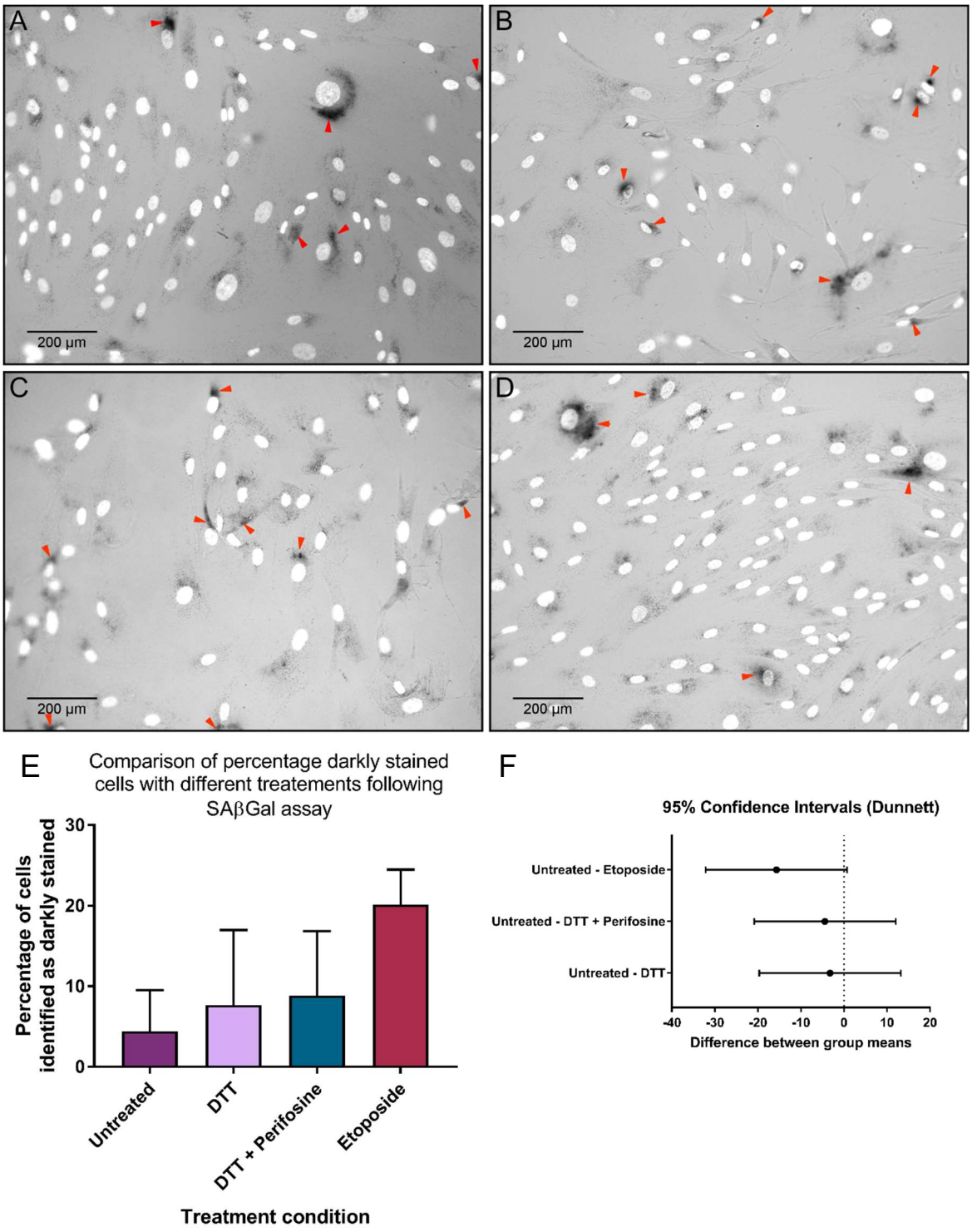


Figure 4.16: Semi-quantitation of SABGal assay. BJ fibroblasts were seeded onto coverslips and treated with **A.** 5 mM DTT, **B.** 5 mM DTT and 5 µM perifosine, **C.** 12.5 µM etoposide or **D.** left untreated. After treatment cells were returned to fresh media to recover for 4 days before fixing and staining for SAβGal and DAPI. Representative images from 3 biological replicates were taken using an inverted wide-field fluorescence microscope with phase-contrast filter and analysed in Fiji. The number of darkly stained cells (indicated by red arrows) in each replicate sample was divided by the total number of cells visible to obtain the average percentage of darkly stained cells from each treatment condition (**E**). These were compared using a one-way ANOVA (n=3) to determine no statistically significant difference between means ($F(3,8)=2.863$, $p=0.1041$). A Dunnett's multiple comparison post-hoc test was used which additionally confirmed that there was no significant difference between any treatment and the untreated control (**F**).

In these experiments, nearly all the cells appeared to have some positive staining for SAβGal. It was, however, noted that there were some cells with a much heavier, darker staining pattern than others (denoted by red arrowheads in Figure 4.16) and therefore these could potentially be used as a point of comparison between the different treatment conditions. The total number of cells in each image were counted, as were the number of heavily stained cells and the percentages calculated (Figure 4.16E). Using a one-way ANOVA, it was determined that there was no statistically significant difference between the means of any treatment, and a Dunnett's multiple comparison post-hoc test confirmed this, which suggested that DTT, with or without perifosine, did not cause a significant increase in the number of senescent cells. It was therefore determined that reductive stress induced by DTT treatment is unlikely to result in senescence. This is consistent with the data shown in chapter 3 (see section 3.4), where there was no evidence for an induced SASP amongst DTT treated cells.

4.5 Discussion and Conclusion

This chapter has described the ability of DTT to induce a growth factor-independent signalling response in the phosphorylation of Akt, and the diminished phosphorylation of p44/42, that has not been previously described. The phosphorylation patterns of these signalling molecules appear to be specific responses to redox stress as they are not seen with the redox-neutral ER stress inducer thapsigargin. However, they cannot be attributed to a specific type of redox reagent as the reducing agent thioglycolate causes a response more like that of the oxidising agent hydrogen peroxide than DTT.

The potential consequences of the chronic Akt phosphorylation seen in response to long-term DTT treatment have been investigated by exploring the possibility of the induction of cell senescence (section 4.4). The SAβGal assay was used to compare the number of senescent cells between treatment with DTT and etoposide, a positive control for senescence. However, there was not sufficient evidence to conclude that reductive stress can drive cells into senescence.

Additional markers of senescence may lend themselves more easily to confident quantitation of senescent cells in differing treatment groups. For example, Debacq-Chainiaux et. al., described the use of 5-dodecanoylaminofluorescein di-β-D-galactopyranoside (C_{12}FDG), a fluorogenic β-galactosidase substrate, for the quantitative analysis of SAβGal activity[211]. This permits quantitation of positively stained cells by fluorescence intensity, which could also be used to set a threshold level whereby cells that stain weakly positive, for example due to contact inhibition, could be excluded from analysis.

Senescent cells are known to accumulate in aged tissues and have become an attractive target for therapies directed at improving the condition and appearance of aged skin. The pro-inflammatory cytokines and matrix-remodelling proteases released in SASP lead to breakdown of the ECM and loss of skin structural integrity[212]. However, advantages of this activity have been demonstrated in the healing of cutaneous wounds where the anti-fibrotic properties of the SASP that facilitate ECM remodelling may reduce excessive collagen deposition that occurs in rapidly forming fibrotic scar tissue[213,214].

The activation of the Akt and/or p44/42 pathway by reductants could open a new avenue for the treatment of ageing in dermal fibroblasts and optimising wound healing. It is possible to imagine a situation where application of a bespoke reducing agent to wounded skin may encourage positive signalling outcomes in cells to improve appearance of scars through its anti-fibrotic mechanisms.

Of course, it is also important to consider the role of p44/42 signalling in this picture. Here it was shown that long-term DTT treatment led to the dephosphorylation of these kinases and the downstream effects of this signalling pattern have not been investigated. The Ras/Ref/ERK pathway

in which p44/42 are implicated is known to induce both cell proliferation (survival) and cell death. The dephosphorylation and inactivation of this pathway may then be a protective or aggressive response towards cells potentially damaged by reductive stress.

The following chapter of this thesis will use proteomics to further investigate the response to DTT and PDGF at the protein level. This will give an integrated understanding of the biological consequences of the signalling responses seen here, and how these events may be translated into proteomic changes within the cell. It also provides opportunity to further explore the subtle but distinct differences in cell responses to DTT with and without PDGF stimulation.

Chapter 5: The effect of reductive stress on the dermal fibroblast proteome

5 The effect of reductive stress on the dermal fibroblast proteome

5.1 Introduction

Signalling events that may trigger changes to cellular proteomes often rely on the post-translational modification of effector proteins that may occur very rapidly. These effector proteins trigger changes to the abundances of other proteins by the promotion of gene transcription to increase mRNA copy number, and translation of this mRNA into functional proteins. However, alterations to protein abundance within cells does not occur uniquely by the up and down regulation of genes and mRNA transcription. Rather, the addition of post translational modifications to proteins already resident in the cell can modulate their turnover; for example, ubiquitination targets proteins for degradation reducing the protein copy number per cell[215,216].

The proteome of a cell is dynamic, with each protein having a specific half-life that may range from 10 minutes to many hours[216]. The time taken for 50% of the HeLa cell's proteome to turn over is estimated to be approximately 24 hours[217]. Thus, when investigating proteomic changes in response to stress, for example, a balance must be struck in determining the time-frame to examine: too long a recovery time and the cells will likely have regained equilibrium or died; too short and the cells will not have had sufficient time for changes in protein abundance to take effect.

This chapter describes the proteomic analysis of DTT treated BJ fibroblasts in the presence and absence of PDGF. PDGF was given for 24 hours and DTT was added for the final 6 hours of this treatment period before lysis and proteome analysis. It was expected that 6 hours of DTT treatment would allow protein responses to be detected without inducing cell death, consistent with the observations in chapter 4. DDA was first used to gain a broad overview into the similarities of protein profiles detected in the lysates from different treatments. DIA by SWATH acquisition was then used to provide quantitative proteomic data on the relative abundances of protein identifications and how these may change between treatments. An analysis of the variance in protein quantification by SWATH acquisition will be discussed first before evaluating the changes seen with each treatment condition. The GO annotations associated with significantly changing proteins will also be analysed for overrepresentation or enrichment of terms in the data set. Finally, the effect of DTT on the extracellular matrix will be further interrogated.

5.2 Experimental set-up for the proteomic analysis of cellular response to PDGF and DTT treatment

BJ fibroblasts were treated for 24 hours serum free media either alone (untreated); with 5 mM DTT for the last 6 hours (DTT only); with 10 ng ml^{-1} PDGF-bb in serum free media (PDGF only); or with PDGF-bb in the serum free media and 5 mM DTT for the last 6 hours (PDGF and DTT). Following treatments, three biological replicates of each condition were lysed in RIPA buffer and prepared for mass spectrometry analysis using FASP. For SWATH acquisition, three technical replicate runs were undertaken for each biological replicate providing a total of 9 replicates for each condition (Figure 5.1). A total of 4487 proteins were successfully quantified using SWATH acquisition; 45 showed a significant (FDR adjusted $p < 0.05$) fold change (FC) greater than 2 with DTT treatment alone, and 28 showed a significant change with DTT and PDGF, whilst only 10 met these criteria when comparing PDGF to Untreated samples.

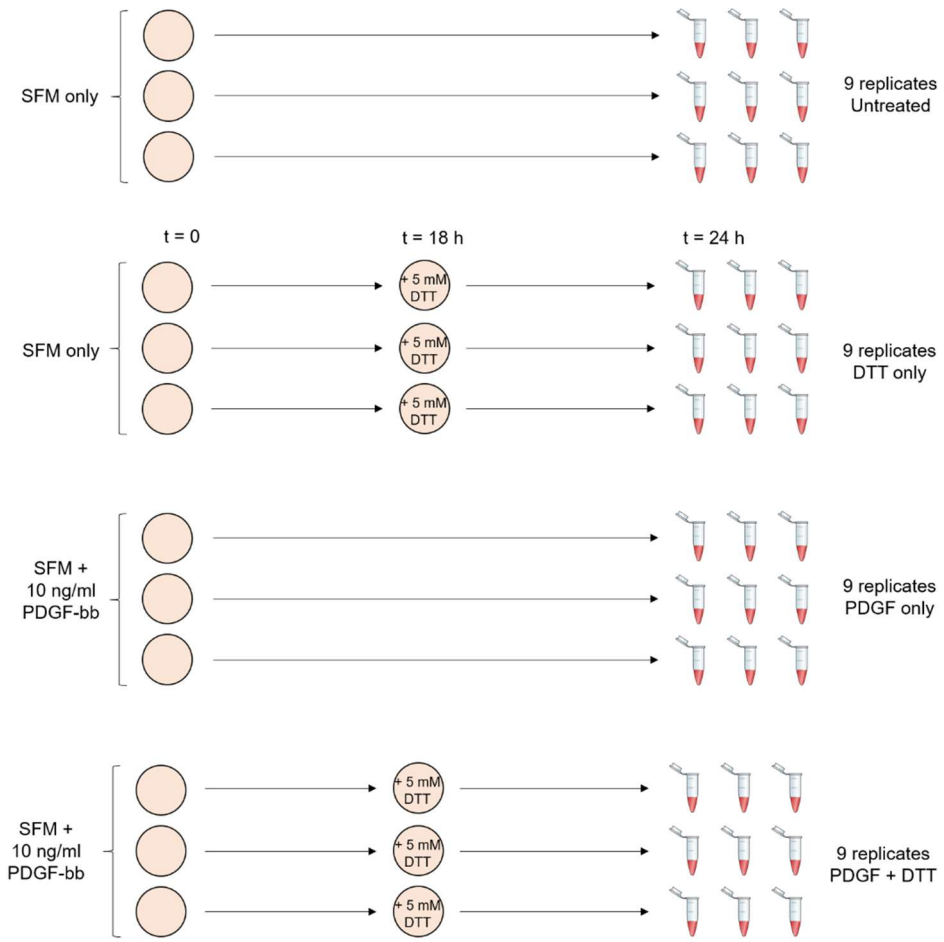


Figure 5.1: Schematic of treatment for proteomic analysis. BJ fibroblasts were treated with serum free media with or without 10 ng ml^{-1} PDGF-bb for 24 hours, with or without 5mM DTT added 6 hours before the end of treatment. Three biological replicates were made for treatments, each of which was used for three technical replicates in MS analysis.

5.2.1 DDA reveals unique protein profiles for each treatment condition

Prior to quantitative proteomic analysis, lysates were subject to DDA to profile the proteins detected in each condition. Across the four treatments, a total of 2243 proteins were identified using this technique. Comparing the identifications demonstrated that, while a substantial proportion of the proteins were the same across all categories, there was a subset of unique proteins seen in each treatment (Figure 5.2). As expected, proteins identified in the DTT treated categories included chaperone-associated proteins such as members of the DNAJ superfamily that are involved in the stimulation of Hsp70 ATPase activity to mediate its binding to unfolded peptides. Stress inducible members of the DNAJ superfamily of proteins have been identified[218]. Therefore, it is plausible that DTT treatment is inducing the upregulation of these proteins as part of a response to ER stress.

Whilst DDA data can indicate which proteins may be involved in responses to treatment, without quantitation it is difficult to determine whether there is a real difference. Precursor ions only fragmented for protein identification in the DTT treated lysate may correspond to a protein genuinely upregulated in response to treatment. However, they may have been selected for fragmentation in only the DTT treated sample due to variation in MS1 precursor ion selection between samples. Therefore, DIA was employed to provide relative quantitation of proteins identified in the lysates from the four samples. DIA by SWATH acquisition fragments all peptide precursor ions and so it is possible to compare the relative abundance of proteins across the samples from different treatments.

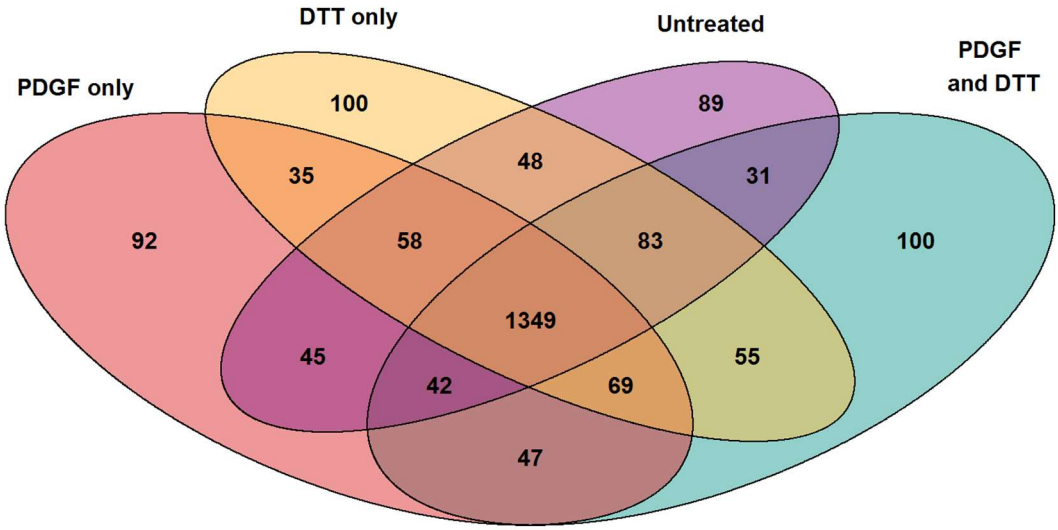


Figure 5.2: Comparison of DDA protein identifications between treatment groups. Venn diagram displaying the number of DDA protein identifications in lysates from each treatment condition, and their overlap.

5.2.2 Microflow SWATH acquisition is more stable than nanoflow

There are two options for SWATH acquisition with SCIEX TripleTOF mass spectrometers: nanoflow or microflow. In nanoflow mode, an ionisation source is used that permits very low flow rates from the liquid chromatography (LC) component of 300 nl min^{-1}. This mode generally has increased sensitivity, and only requires 1 µl sample to be injected which is beneficial in cases were sample availability is low. Microflow mode uses a duospray ion source that permits ionisation of peptides by both electrospray ionisation (ESI) and atmospheric pressure chemical ionisation (APCI) and thus allows complete separation of sample and calibrant material. The peptide sample is ionised using ESI while the calibrant injections are ionised using APCI. This means that there is no contamination of sample from calibration injections that occur between samples in a large batch, for example[219]. The microflow running mode for SWATH acquisition permits flow rates of down to 5 µl min^{-1} requiring a larger injection volume of 4 µl. This slightly reduces the sensitivity of the identifications possible, but the runtime is reduced and stability of retention times increased lending itself to more robust quantitative data sets[220].

When comparing the total ion chromatograms from nanoflow and microflow injections (Figure 5.3A and B respectively), decreased variation in the traces was also seen with microflow injections. This variation does not directly affect quantitation, as intensities are normalised for total peak areas during analysis, but the more variability there is the more adjustment will need to be made which could affect the precision of results.

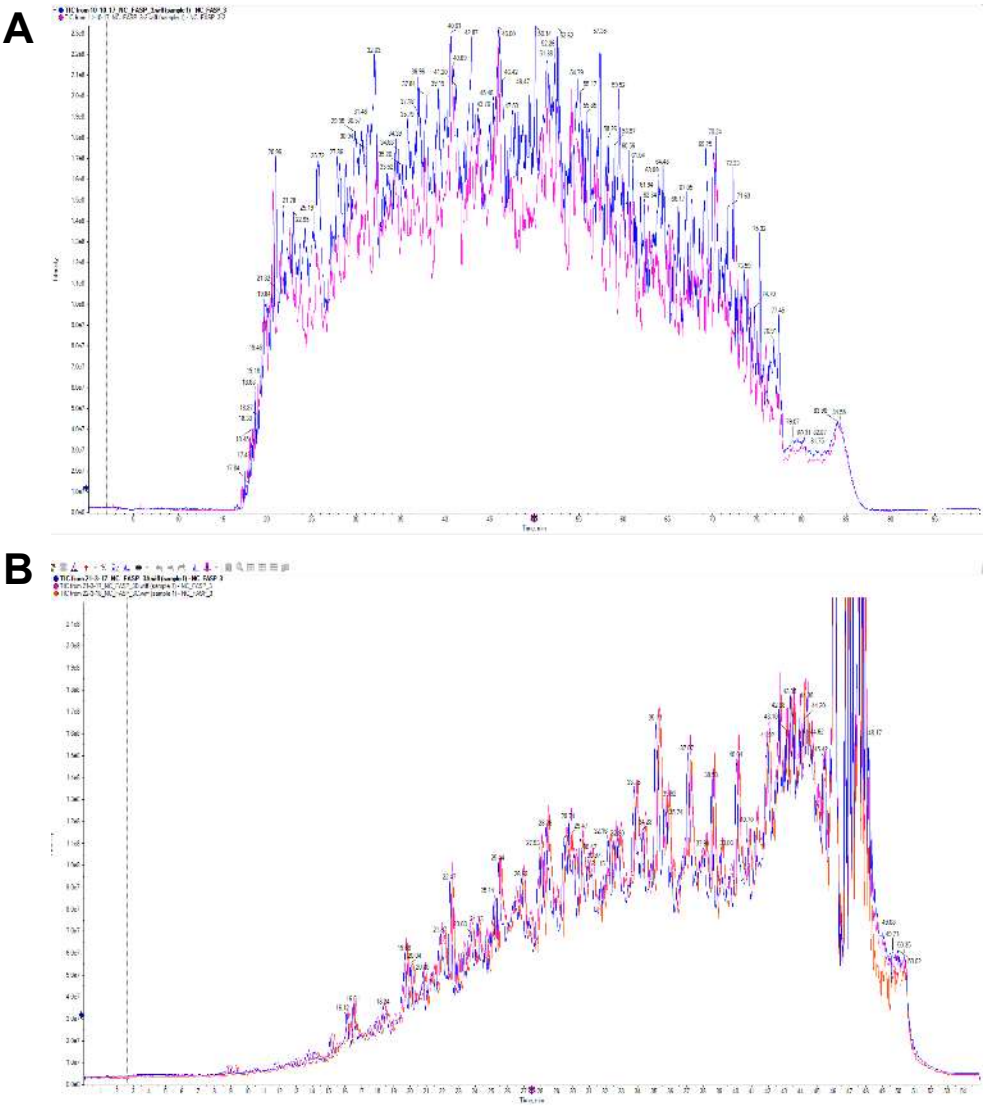


Figure 5.3: Total ion chromatograms from Micro flow SWATH acquisition overlay more tightly than Nano flow. Total Ion chromatograms from technical replicates of a single sample from **A.** Nano flow and **B.** micro flow sources were overlaid to assess stability and reproducibility of ion intensities. TICs overlaid more completely in results obtained from microflow than nanoflow.

A comparison of the results from the two flow sources demonstrated that 495 more proteins were identified with nanoflow alone than microflow alone, although there were unique identifications in both cases (Figure 5.4A). Comparing the number of peptides identified per protein in the two methods demonstrated only minimal differences: while nanoflow showed a slightly higher proportion of proteins identified by more peptides (Figure 5.4B), microflow has a smaller range of values, as well as proteins identified with the maximum number of peptides seen (Figure 5.4C).

As there was sufficient protein in the samples, it was decided to run all the injections in microflow mode due to its increased stability and decreased run times in comparison to nanoflow. There were not sufficient differences in the protein identifications to suggest that nanoflow was better after accounting for the higher levels of variability known to occur in retention time and TICs. It is interesting to note, however, the differences in protein identification and the possibility that a more complete data set for proteome profiling may be obtained by combining data from both flow types.

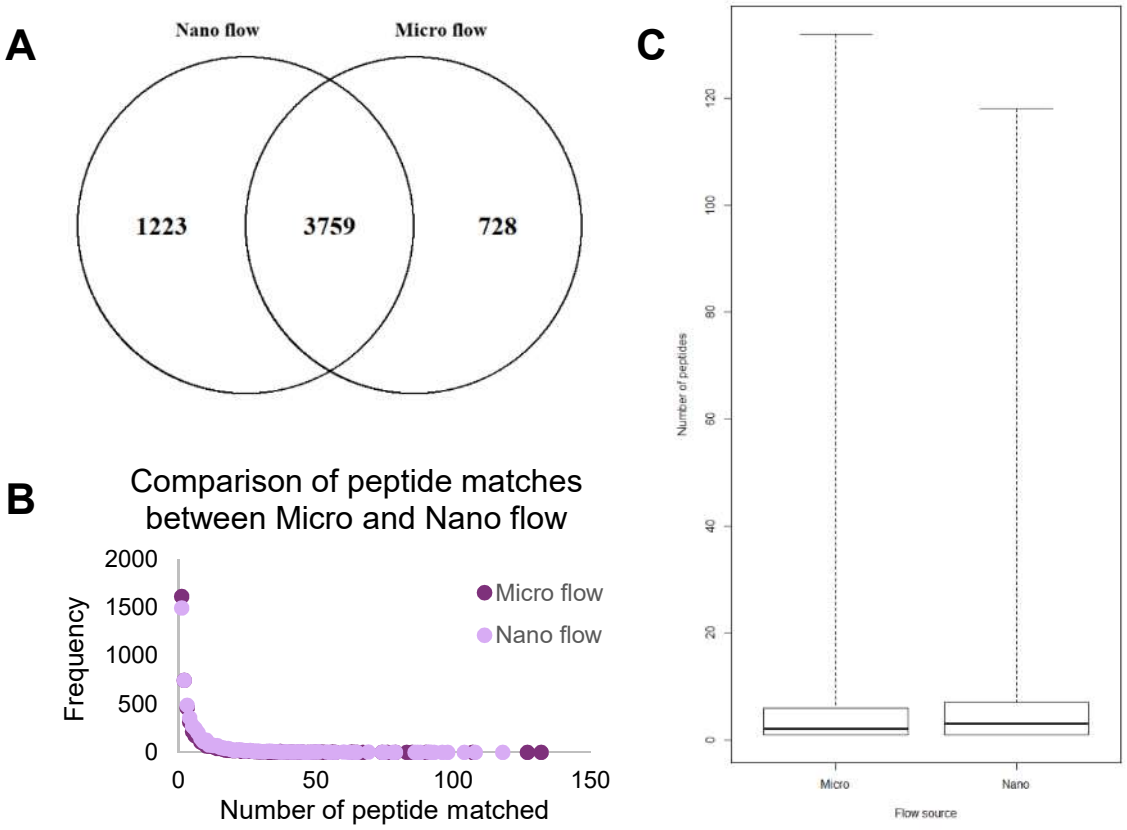


Figure 5.4 : Comparison of results from Nano and Micro flow sources. A. Venn diagram comparing the number of proteins identified in Nano and Micro flow acquisition modes. **B.** Comparison of the number of proteins with varying numbers of peptide matches. **C.** Box plot to compare variation in distributions of the number of peptide matches between Nano and Micro flow.

5.2.3 Analysis of variance amongst replicates

Both biological and technical replicates were used to control for reproducibility of protein identifications within the samples. The average peak area across all 9 result sets was used in comparisons but analysis into the variation between replicate samples was also performed to establish if, for example, there was a replicate lying outside the trend for the data set. To compare the variance within samples, the coefficient of variance (CV) was used. The CV provides a measure of the variance as a percentage of the mean, which gives an indication of how relevant the variation is to the sample. For example, a difference of 1 away from a mean of 100 is much less significant than a difference of 1 away from a mean of 10 however, the standard deviation does not describe this. Dividing the standard deviation by the mean of the data gives the percentage CV, which will be 10-fold greater for a difference of 1 away from a mean of 10 than from a mean of 100. Generally, a %CV value of up to 20% is accepted for technical replicates, and up to 30% in biological replicates (personal communication with Hubbard group, Manchester University).

Therefore, the %CV was calculated for the technical and biological replicates for all identifications with FDR < 5% (Figure 5.5). This demonstrated that most of the peptides were within the accepted thresholds for CV. Within the DTT technical and total replicates analysis, 85 and 86% of the identifications were within the CV thresholds respectively. For the untreated samples this increased to 93 and 97%. The increased variance in the treated sample across all replicates was expected because the response of cells to treatment will inevitably be more variable than cells in their natural state. Even within tissue *in vivo* it is unlikely that cells would respond uniformly to an environmental change. The percentage of identifications that were within the CV threshold in the analysis of technical replicates for PDGF and DTT treated and the PDGF only treated samples were 86 and 88% respectively. The same values were seen for the analysis of all replicates, including biological, together.

Another, and arguably more accurate, way of analysing the proportional variances with respect to the mean is to log transform the peak area values. The logarithmic scale separates values proportionately such that a doubling event and a halving event are both given the same absolute value. In a scale that is not logarithmic, if a value of 10 doubled, the result would be 20, whereas if it halved the result would be 5. Proportionately both are a 2-fold change, but the changes (i.e. resulting value minus original value) are listed as 10 and -5 respectively, which are very different. On a logarithmic scale log(20) = 1.3, log(10) = 1, and log(5) = 0.7 and therefore the changes for the doubling and halving events are 0.3 and -0.3 respectively, both the same absolute value. The log transformation also means that the variances, and therefore standard deviations, are proportional to the mean. Looking at the data from the mass spectrometers, it is possible to use log transformations to measure the proportional standard deviation of the replicate data (Figure 5.6).

This comparison of the standard deviations of log values demonstrates a difference in the variance of PDGF treated samples compared to that of the DTT only or untreated samples. This difference is not as obvious in the CV comparisons and may provide an explanation as to why there are fewer significantly changing proteins in the PDGF treated samples, as increased variation will increase the p-value and so there may be truly changing proteins that are not reaching the threshold of significance.

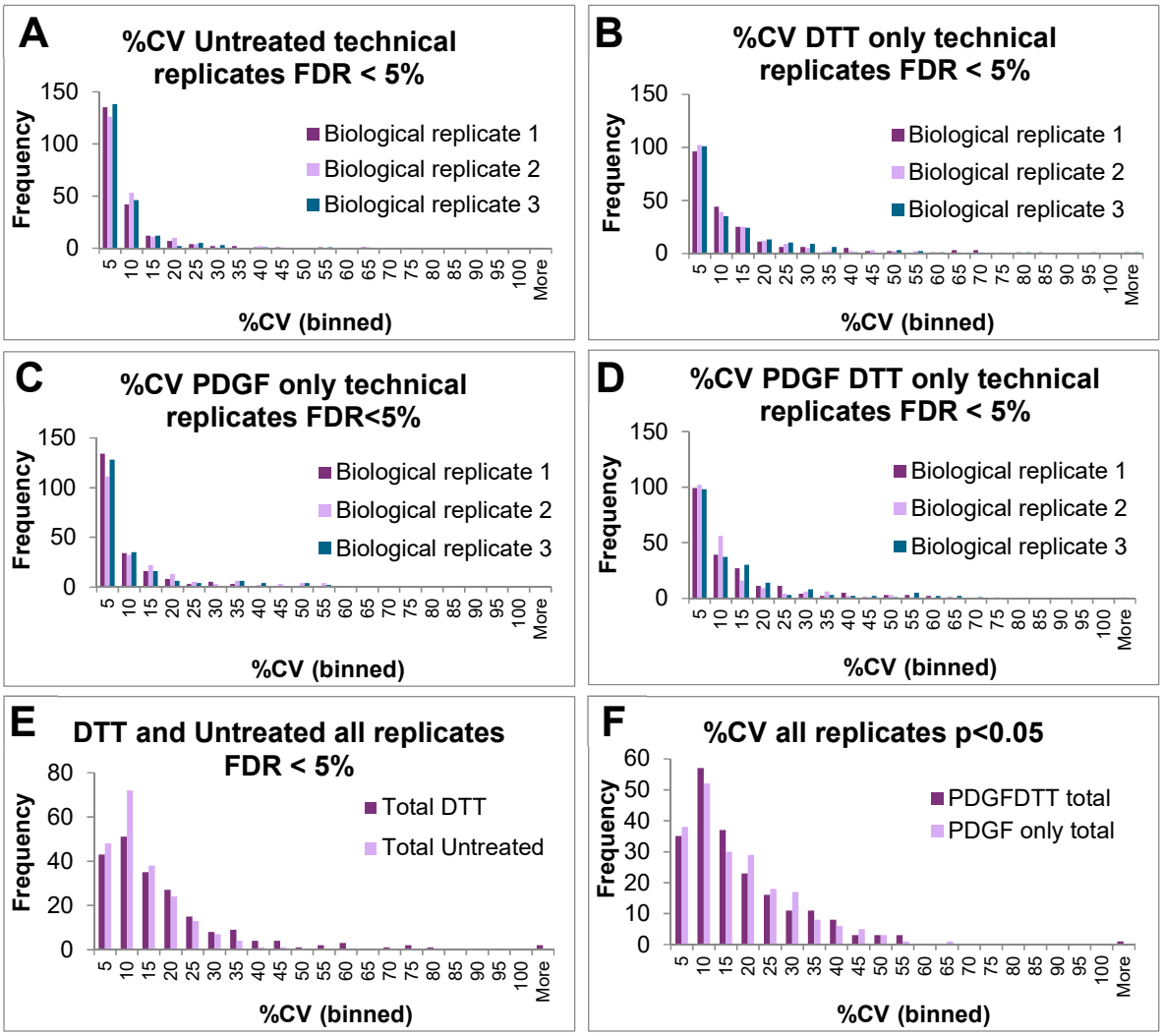


Figure 5.5: Histograms display distribution of coefficient of variance values across replicates. A-D show coefficient of variance values for technical replicates in each of three biological replicates. **E** and **F** show coefficient of variance values for all replicates, both technical and biological, combined.

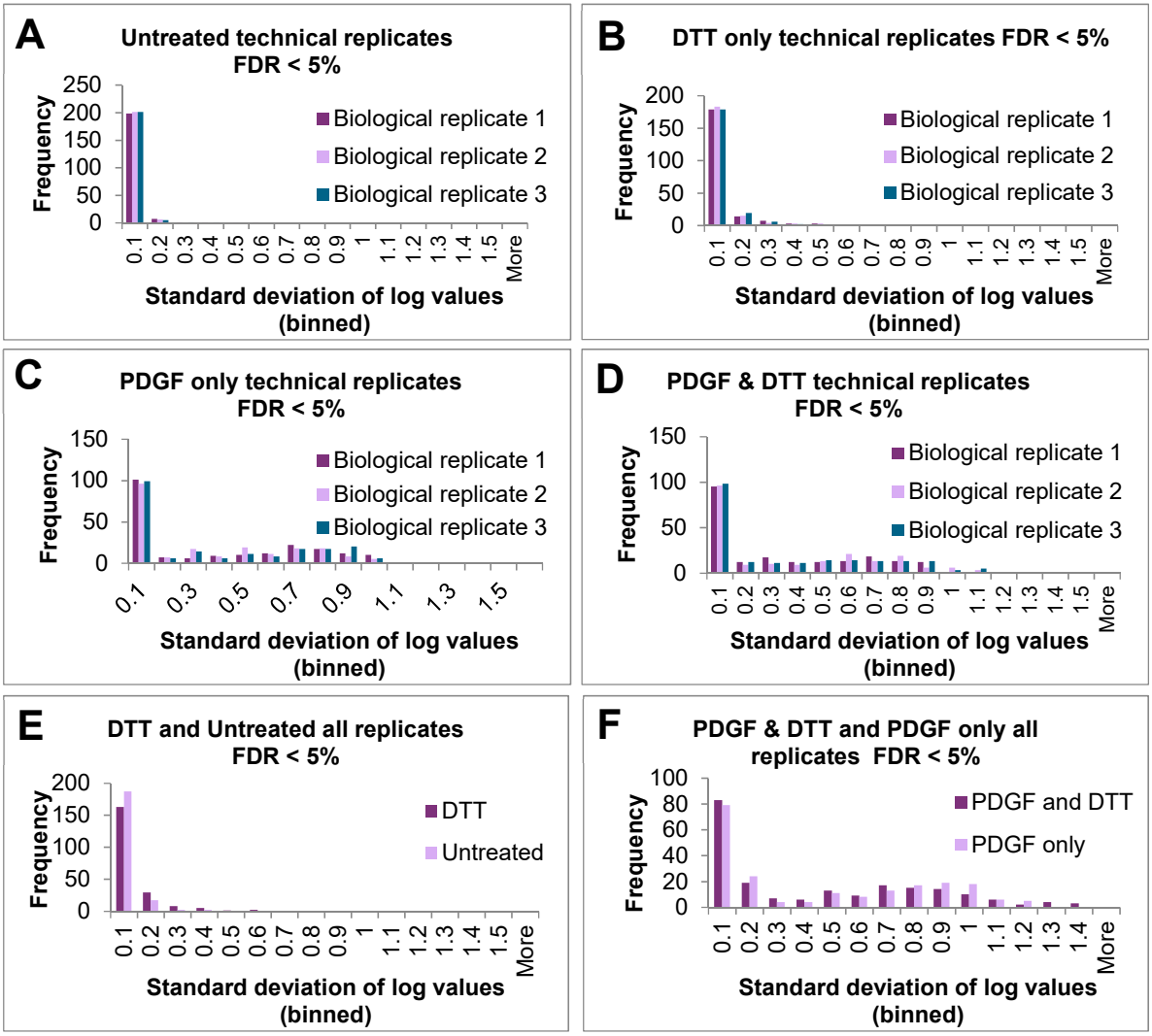


Figure 5.6: Histograms display distribution of standard deviation of log values. A – D show standard deviation of log values for technical replicates in each of three biological replicates. **E** and **F** show standard deviation of log values for all replicates, both technical and biological, combined.

5.3 Protein responses to PDGF and DTT treatments

Proteins were determined as significantly changing when their fold change was greater than 2 (doubling event), or less than 0.5 (halving event) and the FDR adjusted p-value from a Student's t-test was less than 0.05. To determine the spread of differential quantifications the log fold change was plotted against the negative log of FDR adjusted p-value in each of three comparisons (PDGF to untreated; DTT to untreated; and PDGF and DTT to PDGF only) (Figure 5.7, Figure 5.8, and Figure 5.9 respectively).

As can be seen from these plots the response is predominantly a down-regulation of proteins in response to DTT treatment both with and without PDGF. The response to PDGF alone appears more balanced, although there are substantially fewer proteins showing a significant response in this comparison. It is possible that with a 24-hour stimulation, as was given in these experiments, the cells are recovering from their initial response by the time of lysis and therefore protein changes are more difficult to detect at a fold change of more than 2.

Histone H4 (HIST1H4A) was shown to be decreased in response to PDGF (Fold change=0.41; adjusted p-value = 5.55×10^{-5}) in this data set. Previously, PDGF-bb has been shown to reduce the expression of structural maintenance of chromosomes (SMC) genes, which was mediated by a variety of mechanisms including the deacetylation of histones[221]. Histone deacetylation induces closer interactions between histone proteins and DNA, making the DNA less accessible to transcriptional machinery and therefore represses gene expression. DNA, and any tightly associated proteins, is removed by centrifugation following cell lysis. It is possible that, if PDGF stimulated histone deacetylation of some genes, the histones now tightly associated with DNA were not identified in our proteomic analysis leading to an apparent reduction in histone proteins following PDGF treatment.

Pre-laminA/C (LMNA), which is cleaved into Lamin A/C, was identified as being significantly upregulated in response to PDGF. The Lamin gene contains an alternative 5' splice site (SS) termed the "progerin 5'SS" because the alternative splicing of the lamin gene by this SS leads to the excision of intron 11 and production of the defective progerin protein[222]. In healthy cells, this occurs only at a very low level but in Hutchinson Gilford Progeria Syndrome (HGPS) the progerin 5'SS is activated leading to the accumulation of progerin and defective nuclear envelope morphology[222]. Vautrot et al., demonstrated that treatment with PDGF-bb led to decreased usage of this SS and increased accumulation of correct lamin mRNA and protein within HGPS cells[222]. It is possible that, in healthy cells, PDGF is still able to promote correct splicing of lamin to increase its expression. Alternatively, the increase seen may be a simple product of increased cell proliferation leading to upregulation of nuclear structural proteins in order to support the nuclear division required for mitosis. This theory may also explain the increase in cytospin-B (SPECC1), another nuclear structure protein.

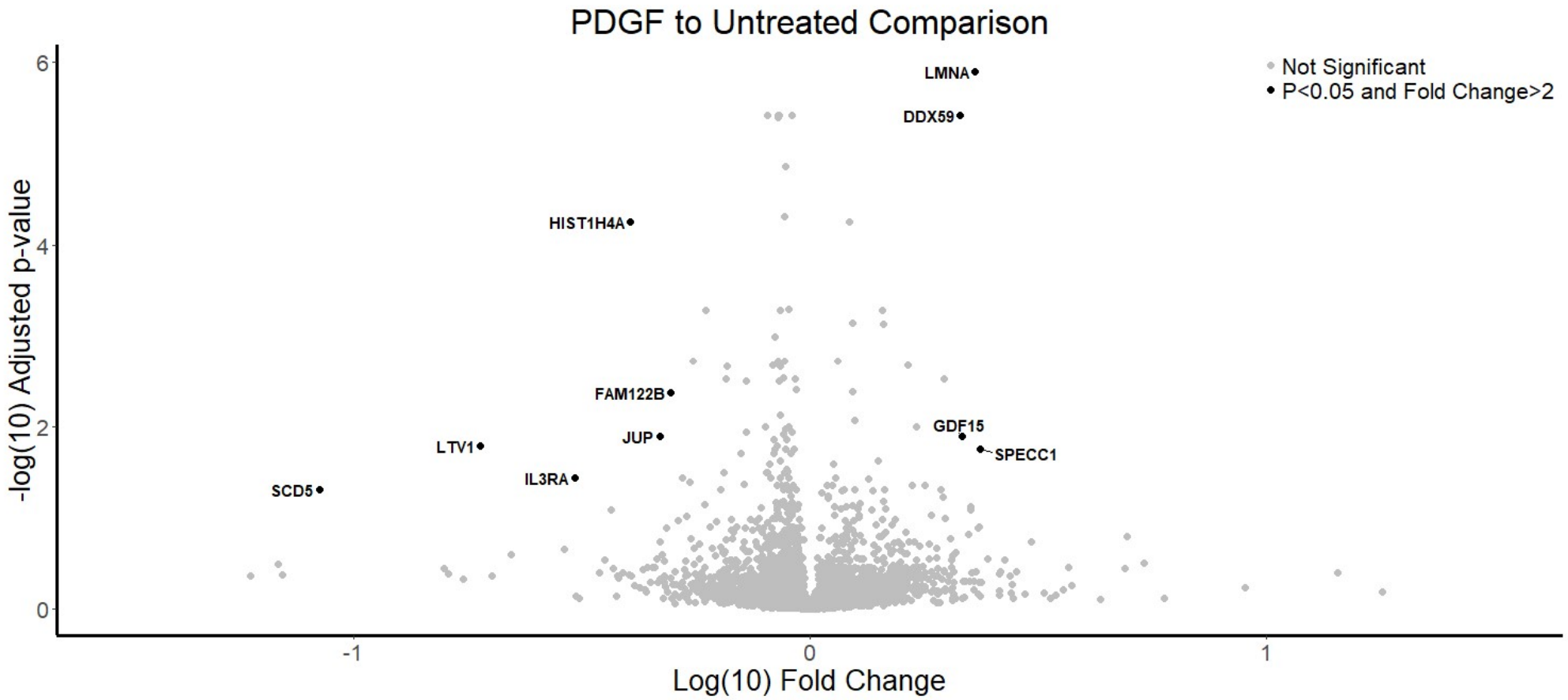


Figure 5.7: Volcano plots display distribution of significantly changing proteins across fold change and significance values (part A). Log10Fold change is plotted against the negative log10FDR-adjusted p-value for proteins quantified in the comparison of PDGF to Untreated samples.

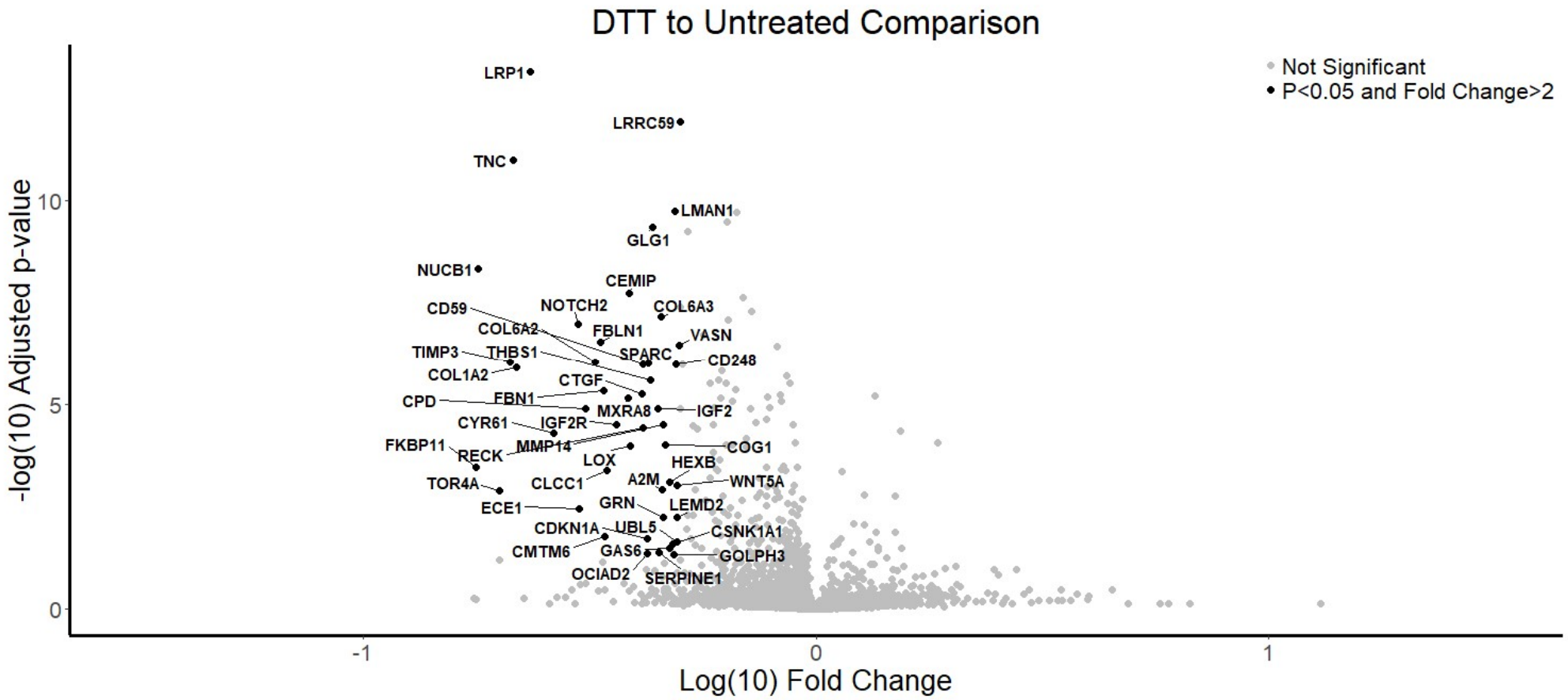


Figure 5.8: Volcano plots display distribution of significantly changing proteins across fold change and significance values (part B). Log10Fold change is plotted against the negative log10FDR-adjusted p-value for proteins quantified in the comparison of DTT to Untreated samples.

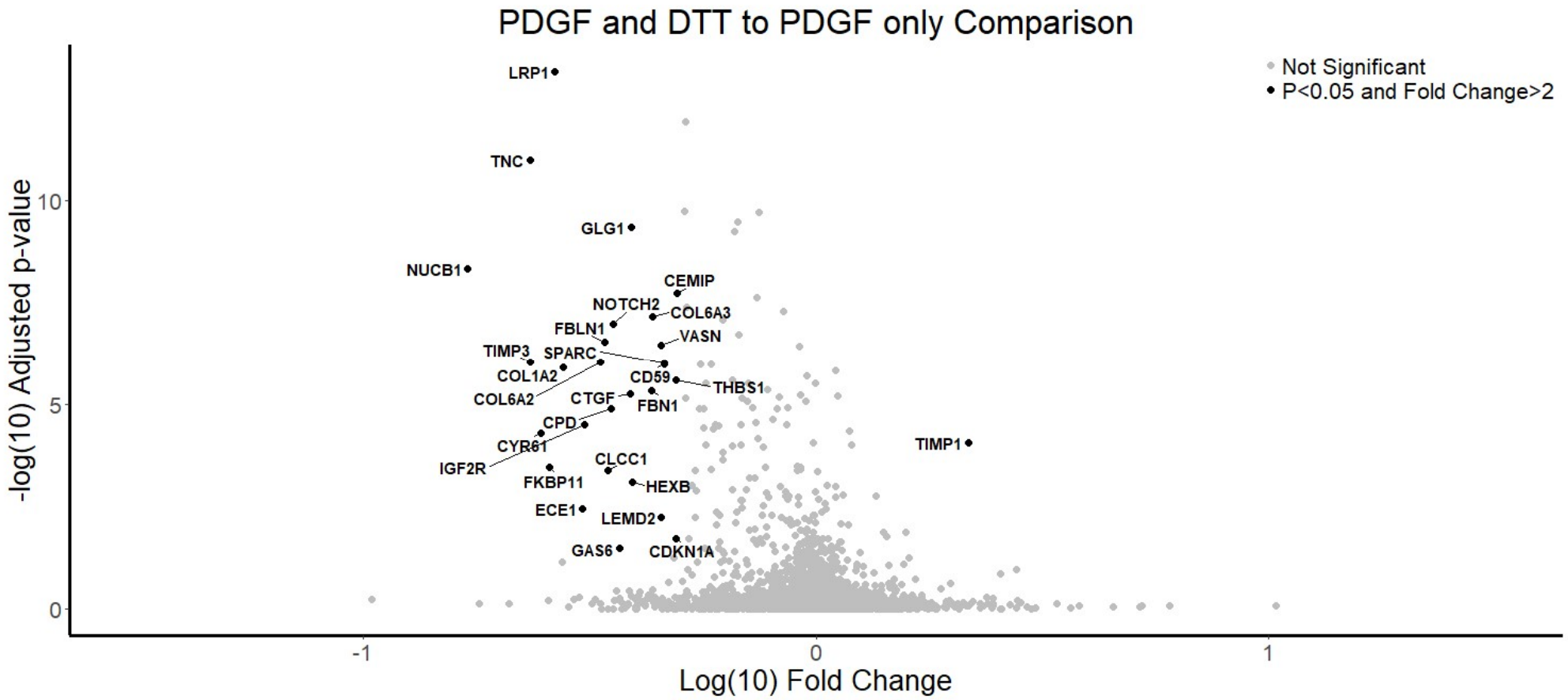


Figure 5.9: Volcano plots display distribution of significantly changing proteins across fold change and significance values (part C). Log10Fold change is plotted against the negative log10FDR-adjusted p-value for proteins quantified in the comparison of PDGF and DTT to PDGF only samples.

Comparing those proteins that were identified as significantly changing in DTT treated samples with and without PDGF, it became obvious that there was a significant overlap in the two responses. Almost all the significantly changing proteins in the PDGF treated comparison were also identified in the DTT only to untreated comparison, however there were an additional set of proteins only seen when PDGF treatment was absent (Figure 5.10). The single protein that is unique to the PDGF treated comparison is the only upregulated protein seen: TIMP1, which demonstrated a fold change of 2.17 and adjusted p-value of 8.62×10^{-5}. TIMP1 inhibits the proteolysis of ECM components and it has also been suggested that TIMP1 is involved in the reduction of inflammation at wound sites[223]. PDGF-bb has been previously shown to increase TIMP1 production in fibroblasts[224,225], implicating TIMP1 in the regulation of ECM re-modelling in response to mitogenic signals and this is therefore an expected response. It is interesting to note the opposing down regulation of TIMP3 (fold change: 0.23; adjusted p-value: 9.25×10^{-7}). TIMP3 is uniquely involved in the inhibition of ADAM17, a protease which cleaves the membrane bound tumour necrosis factor alpha (TNFα) into its soluble form. Release of TNFα into the ECM triggers inflammation, and it is therefore suggested that TIMP3 is a key regulator of this effect[226]. It seems contradictory that there is an upregulation of TIMP1 that may play a role in reducing inflammation while TIMP3, whose absence may lead to increased release of the pro-inflammatory cytokine TNFα is reduced. It has been suggested in the literature that the TIMP family members may have opposing effects on cell survival with TIMP1 and 2 indicated as suppressors of apoptosis while TIMP3 and 4 promote apoptosis in a cell-type specific manner (reviewed by Gardener and Ghorpade[227]). Thus, it is possible to theorise that the apparently opposing responses of TIMP1 and TIMP3 seen in PDGF and DTT treated fibroblasts here, may be part of a wider response from the cells to promote cell survival.

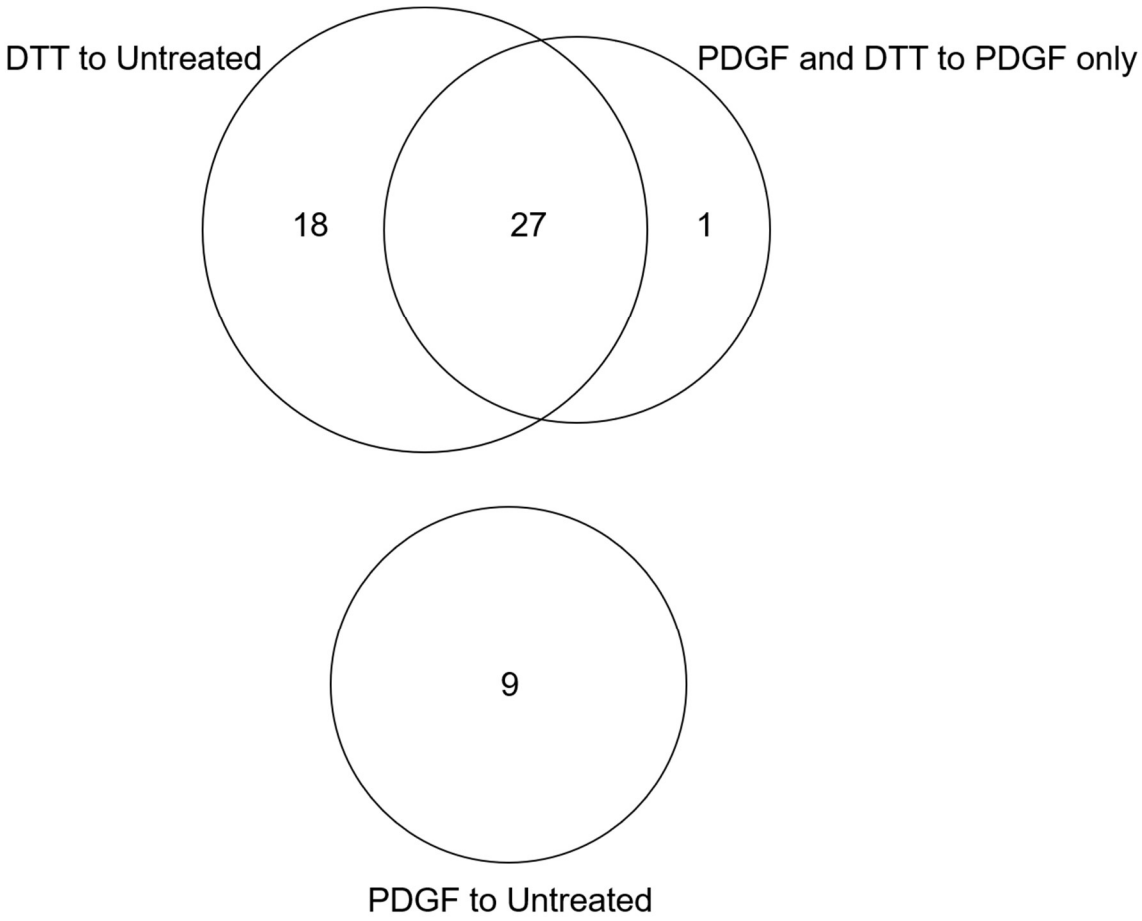


Figure 5.10: The response of BJ fibroblasts to DTT is broadly similar in the presence and absence of PDGF, and distinct from the response to PDGF alone. The proteins identified as significantly changing with DTT treatment were compared between treatments with and without PDGF, and to those identified as significantly changing in response to PDGF alone. This demonstrated that fewer proteins responded to DTT in the presence of PDGF, and that the proteomic response to PDGF alone was less than, and completely distinct from, the proteomic response to DTT.

The downregulated proteins (+ DTT, ± PDGF) were directly compared (Figure 5.11), with all of them centring around a 2-fold change (≤0.5). The changes seen in the PDGF treated cells were slightly different to those seen in the cells treated without PDGF, often with fold changes slightly diminished. However, comparing these by a Student's t-test determined that the only difference that was statistically significant was that seen in COL6A3 where the fold change with DTT alone is 0.456 but 0.436 in the presence of PDGF. This indicates a slightly greater decrease in COL6A3 in response to DTT with PDGF than without.

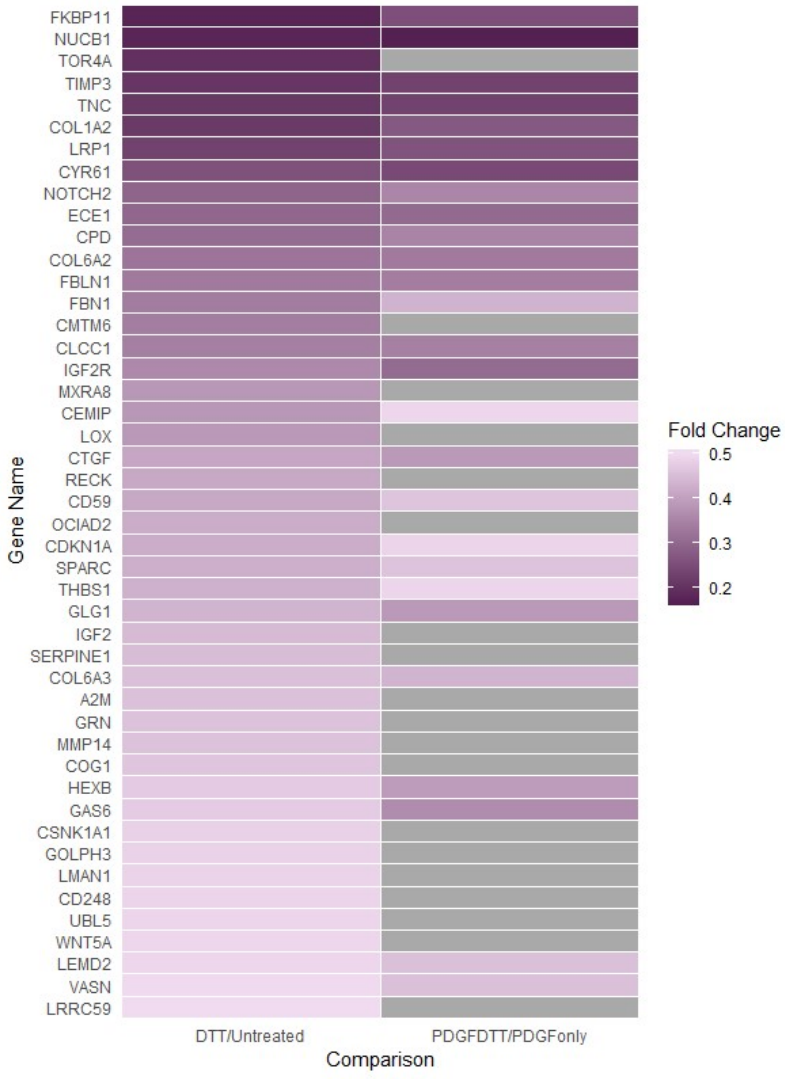


Figure 5.11: Fold changes appear slightly diminished with PDGF than without. The fold change of proteins identified to be significantly reduced following DTT treatment is plotted on heatmap for DTT to Untreated comparison and PDGF and DTT to PDGF only comparison. Proteins which were not significantly reduced in the presence of PDGF are coloured grey.

5.3.1 Statistical enrichment test reveals enrichment of biological adhesion in response to DTT

The statistical enrichment test determines whether the distribution of fold change values for a GO term is significantly different from the overall distribution of all values using a Mann Whitney U test. Using the complete GO biological process annotation there was no statistically significant enrichment seen in either comparison (DTT to Untreated, or PDGF and DTT to PDGF only). However, using the PANTHER GO Slim annotation for biological process, a significant enrichment was seen for the term 'biological adhesion' in the DTT to Untreated comparison only (Figure 5.12).

PANTHER GO Slim is a reduced version of the full GO database where only the main terms are used, reducing the specificity and complexity of the annotations, thus providing a broader overview of the content of data sets. These broader GO Slim terms will therefore contain more annotated proteins than each of the specific terms from the full GO and this is likely why statistically significant enrichment is only seen with this database. With a term that only holds a few protein annotations, the distribution of values would have to be much more specific to be identified as significantly different from the overall distribution of all terms. In the GO Slim database terms have a higher number of annotated proteins, and distributions are therefore more able to withstand some variances while still demonstrating statistically significant differences to the overall pattern of terms.

'Biological adhesion' is defined as "the attachment of a cell or organism to a substrate or other organism" and was annotated to 119 proteins in this data set including, for example, ECM proteins such as collagens, cell junction proteins including integrins and cadherins, as well as signalling proteins such as the Ras related protein RAP1A. The enrichment shows that the distribution of proteins annotated to this term is shifted more towards the lower values than the overall distribution pattern of all proteins. This suggests that proteins associated with 'biological adhesion' are disproportionately downregulated following DTT treatment compared to the global proteome response. Anoikis, a form of apoptosis, is triggered in cells by insufficient cell-matrix attachments[228]. A downregulation in integrin proteins, which are transmembrane proteins that play a fundamental role in the attachment of the intracellular cytoskeleton to the ECM, may indicate a reduction in possible attachment points of the cell. Equally, a loss of collagen proteins (also annotated here as belonging to the subset of proteins labelled with the term "biological adhesion") may also lead to the breakage of integrin/collagen linkages and prevent sufficient cell-matrix attachments from being maintained.

The integrin proteins were not seen in the list of proteins with a statistically significant response (Figure 5.11), however re-examining the data for these proteins showed fold changes of just below the threshold of 2 or 0.5. Integrins $\alpha 2$, $\alpha 5$ and $\beta 1$ showed a fold change of 0.66 ($p=0.012$), 0.64 ($p=3.47\times10^{-10}$), and 0.61 ($p=7.19\times10^{-6}$) respectively. These fold changes suggest a change in protein abundance of approximately 1.5 times in the negative direction. It will be important to consider the possible exclusion of proteins of interest by the stringent parameters set in future analysis.

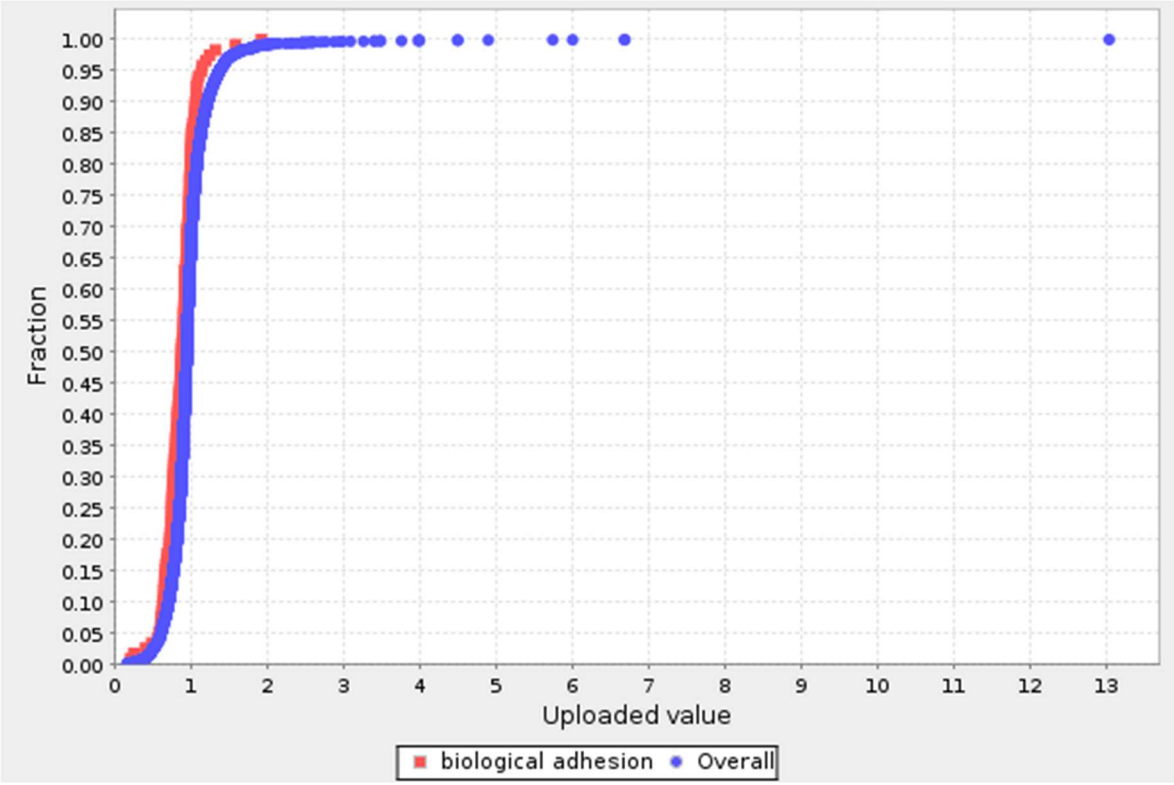


Figure 5.12: Biological adhesion is significantly enriched in DTT to Untreated comparison results. Cumulative fraction plotted against uploaded value of fold change for all proteins quantified in DTT to untreated comparison. Distribution of overall values shown in blue, and distribution of values annotated to biological adhesion shown in red. PANTHER Enrichment Test (Released 20170413) was used with PANTHER version 13.1 Released 2018-02-03 and PANTHER Go-Slim biological process annotation data set. Bonferroni correction for multiple testing used for significance values. Biological adhesion found to be significantly enriched with p-value of 2.33×10^{-4}.

5.3.2 Statistical overrepresentation testing suggests disproportionate response to ECM related proteins

The GO terms annotated to the proteins identified as significantly changing in response to DTT treatment were next tested for statistical overrepresentation using the Panther overrepresentation test[143,166] (Figure 5.13). This uses a Fisher's exact test with FDR correction for multiple testing to determine whether the proportion of results in the input annotated to a specific GO term is significantly different to the proportion of results in a reference list annotated to the same term. Thus, it provides an idea as to which terms have more than the expected number of genes annotated to them based on a reference data set. The significantly changing proteins were tested against all proteins identified in the lysates. Overrepresented terms were divided into three categories: those that were only overrepresented in PDGF treated results, those that were only overrepresented in results where PDGF was absent, and those that were overrepresented in both conditions (±PDGF).

Due to the hierarchical nature of the GO database there is varying specificity in the terms. For example, CYR61 is annotated to 'regulation of cell migration' in +PDGF but 'cell migration' in -PDGF. The higher level the term in the GO, the more annotations it holds. It is probable, therefore, that a higher number of proteins annotated to 'cell motility' (a high-level term) are required for overrepresentation to be significant than are required for 'regulation of cell migration' (a lower level term). As expected, 8 terms were annotated to 'regulation of cell migration' in the presence of PDGF, while 11 terms were annotated to 'cell migration' in the absence of PDGF. As all proteins, except for TIMP1, were downregulated in response to DTT, overrepresentation here indicates that the number of proteins associated with a given term is more than would be expected if downregulation was uniform and proportional across all proteins identified. It therefore suggests that cell migration is disproportionally down regulated in response to DTT. With fewer proteins associated with cell migration in the presence of PDGF, it is possible that the migration promoting effects of PDGF treatment offer the cell some protection from the downregulation induced by DTT treatment. This would therefore reduce the number of proteins associated with cell migration in the PDGF treated samples such that overrepresentation of the broader term 'cell migration' is not statistically significant.

The hierarchical nature of the GO database also means that a single protein is annotated with several terms. As such, there are annotations that are deemed overrepresented and seem out of place in a study of dermal fibroblasts. For example, 'heart development' (+ PDGF) and 'embryonic morphogenesis' (- PDGF) are not terms expected amongst proteins responding to DTT treatment in the skin. However, the proteins associated with these terms are also associated with signalling events and other, more obviously relevant terms.

'Regulation of ERK1 and ERK2 cascade' was overrepresented in the significantly changing proteins from DTT treatment both with and without PDGF (Figure 5.13). This suggests that proteins involved in ERK1/2 regulation are disproportionately downregulated compared to the total quantifiable proteome detected in these studies. This is in concordance with the results from the signalling event analysis discussed in chapter 4 (see section 4.3) in which DTT treatment induced dephosphorylation of ERK1/2. In both cases, with and without PDGF, the proteins annotated to this term were CYR61, NOTCH2, CTGF, TIMP3, GAS6 and FBLN1. CYR61 has been linked to induction of the ERK1/2 cascade in osteosarcoma cells undergoing epithelial to mesenchymal transition (EMT)[229]; inactivation of ERK1/2 inhibited the Jagged/Notch signalling pathway in lens epithelial cells, also in a study of EMT [230]; GAS6 induced ERK1/2 phosphorylation in metastatic prostate cancer cells found in bone[231] and ERK1/2 signalling was shown to upregulate CTGF in mediation of myocardial fibrosis[232]. These findings would all agree with a downregulation of these proteins associated with a decrease in ERK1/2 phosphorylation, in so much as increases in ERK1/2 signalling in the cited literature are associated with increase in the proteins of interest. Thus, conversely, it could be expected that decreases in the protein can be associated with decreased ERK1/2 signalling.

There were no GO terms associated with senescence, which are daughter terms of 'cell ageing' and 'cellular response to stress', seen to be overrepresented in the significantly changing proteins. Two of the proteins in these sets were annotated to senescence: SERPINE1 and CDKN1A. SERPINE1, also known as plasminogen activator inhibitor 1, is a target gene of p53 that was shown to be necessary and sufficient to induce replicative senescence in cells including BJ fibroblasts[233]. Cyclin dependant kinase inhibitor 1, CDKN1A, is controlled by p53 and blocks cell cycle progression by inhibition of cyclin dependant kinases leading to senescence[234,235]. However, as these proteins are decreased in response to DTT, this only confirms that there is no evidence of senescence following DTT treatment in these cells.

Several terms appeared overrepresented in the DTT responding proteins that seemed to be associated with modification of the ECM. These included 'extracellular matrix organisation' overrepresented both with and without PDGF, 'regulation of cell migration' and 'cell migration' as discussed above, and 'collagen catabolic process' overrepresented without PDGF. In addition, the enrichment testing showed that proteins annotated to 'biological adhesion', many of which are ECM proteins, were disproportionately downregulated with DTT treatment (see section 5.3.1). This triggered an interest in the changes that were occurring in ECM proteins secreted by fibroblasts in response to DTT. Fibroblasts are the main producers of collagen to support ECM structure and function and therefore collagen proteins in the complete data set were revisited, to further investigate the changes in response to DTT with and without PDGF.

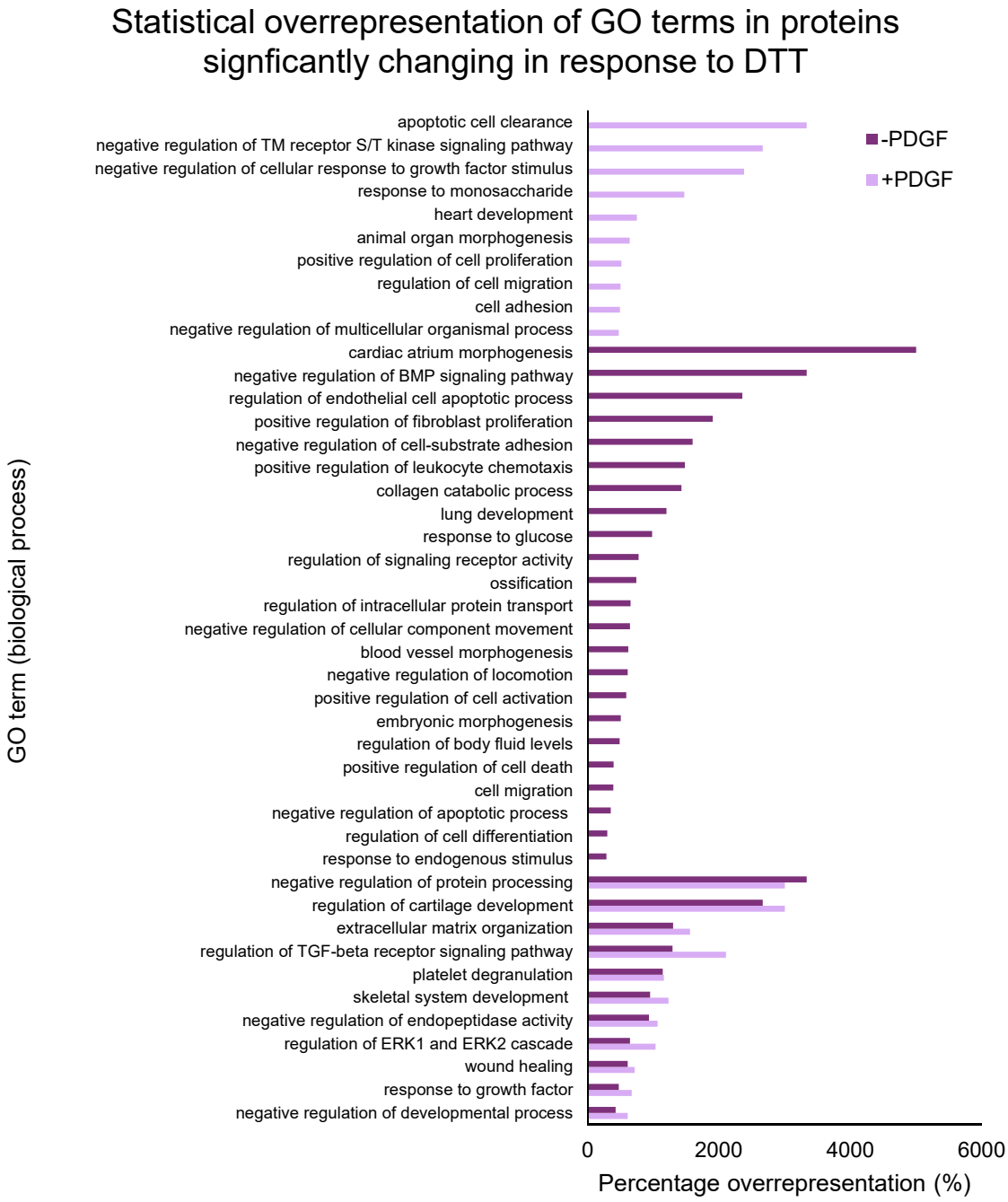


Figure 5.13: Analysis of significantly changing proteins in response to DTT reveals statistically significant overrepresentation of GO terms. Percentage overrepresentation is plotted for GO biological process terms found to be statistically significantly overrepresented in those proteins that significantly changed in response to DTT either with or without PDGF. Significantly changing proteins were tested against a reference list of all quantified proteins in the data set using PANTHER Overrepresentation Test (Released 20171205) and PANTHER version 13.1 Released 2018-02-03. A Fisher's exact test with FDR multiple test correction was carried out using the GO biological process complete annotation set. For ease of viewing, where both parent and daughter terms were found to be overrepresented, only the parent term is displayed.

5.3.3 A set of collagen proteins are diminished following DTT treatment

It was noted that, amongst the 18 collagen proteins identified in the data sets, there were only 3 that met the threshold of 2-fold change. However, a further 6 collagen proteins had a fold change of more than 1.5 (or less than 0.67) with very low adjusted p-values suggesting high levels of significance. They also had many more peptide matches than the 2 usually required to confidently identify proteins from mass spectrometry data (Table 5.1).

Table 5.1: Coverage information for collagen proteins identified as changing in response to DTT. The percentage coverage for collagen proteins with fold change (FC) greater than 1.5 and FDR adjusted p-value less than 0.05 was calculated from peptide matches. Identified peptides used for coverage determination were quantified in all data sets, whether treated or otherwise. Data ranked by percentage coverage.

Gene Name	No. of peptide matches	Percentage coverage (%)	DTT to Untreated		PDGFDTT to PDGF only	
			FC	p-adjusted	FC	p-adjusted
COL1A1	32	32	0.57	2.25E-03	0.68	2.25E-03
COL1A2	30	31	0.22	1.25E-06	0.28	1.25E-06
COL6A3	63	20	0.46	7.05E-08	0.44	7.05E-08
COL12A1	44	18	0.52	5.94E-10	0.66	5.94E-10
COL3A1	14	14	0.65	1.72E-02	0.78	1.72E-02
COL6A2	15	14	0.33	9.38E-07	0.33	9.38E-07
COL6A1	13	12	0.54	1.20E-03	0.65	1.20E-03
COL18A1	7	5	0.61	4.08E-04	0.54	4.08E-04
COL5A2	4	4	0.64	5.23E-03	0.76	5.23E-03

Comparing the percentage change in peak area of the collagens from untreated to DTT treated, either with or without PDGF demonstrated that in nearly all cases, the response of collagens was greater in the absence of PDGF (Figure 5.14). This is consistent with the trend seen previously where the proteins seen to be significantly changing in PDGF treated cells are fewer in number, with fewer statistically significant overrepresented GO terms and no enrichment. It is possible that the growth promoting effects of PDGF treatment provided the cells with some protection against the harmful effect of reductive stress induced by DTT.

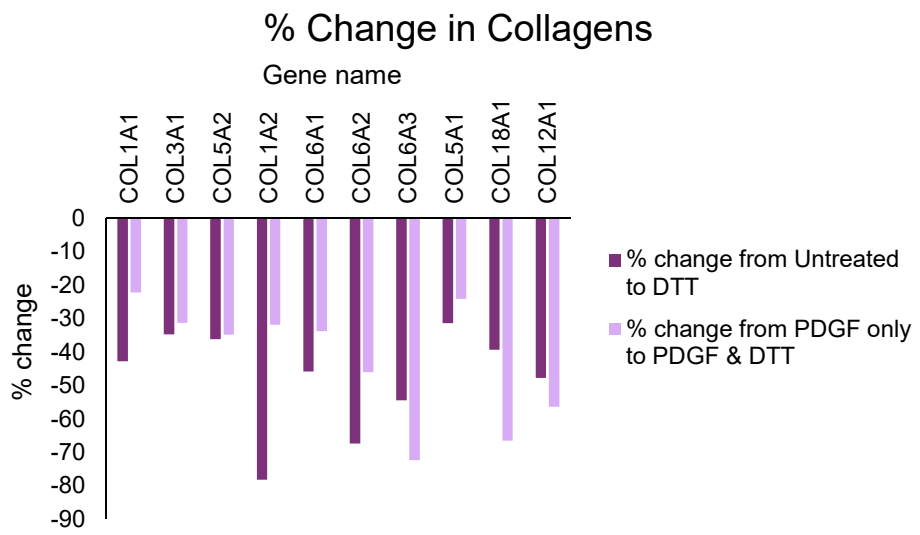


Figure 5.14: Reduction of collagen proteins is enhanced in the absence of PDGF. The percentage change of peak areas quantified for collagen proteins in response to DTT treatment was calculated and compared between conditions with and without PDGF. For most collagen proteins, there was a greater decrease seen without PDGF.

Oxidative stress has been shown to reduce collagen synthesis in rat cardiac fibroblasts[236] however it has previously been demonstrated that renal injury associated with increased collagen production is increased in rats fed on an antioxidant deficient diet that limits peroxide clearance[237]. Whilst research has focussed on the effects of redox agents on collagen in situations such as cardiovascular disease, little is known about the effects of reductants on ECM content in dermal fibroblast cells.

Bleomycin is a drug used in chemotherapy but is associated with the development of pulmonary fibrosis, which may be due to the production of ROS within cells following treatment. In 1989, Wang et. al., demonstrated that treating hamsters with taurine at the same time as bleomycin reduced pulmonary fibrosis including the reduction of bleomycin induced increases in collagen content[197]. Taurine was also shown to inhibit increased collagen production triggered by high levels of glucose in mesangial cells, but not epithelial cells and was thus deemed to be cell-type specific[198]. Although now dated, these investigations suggest a role for redox regulation in the maintenance of collagen production in a physiological setting. Taurine induced signalling responses similar to those seen with DTT treatment (see section 4.3.4) and so this may indicate a role for reductants in the regulation of collagen production in BJ fibroblast cells.

5.4 Discussion and Conclusion

This chapter has described the proteomic responses to DTT treatment seen in BJ fibroblasts with and without PDGF. Variance analysis from proteomic quantification was carried out before determining and investigating significantly changing proteins. The response seen in both cases was broadly similar, although the addition of PDGF decreased both the number of responding proteins and the overrepresentation/enrichment of GO terms. Finally, the effect of DTT on ECM collagens was investigated demonstrating a significant decrease in the amount of collagen that was just below the 2-fold-change threshold used to select responding proteins.

PDGF is known to induce an increase in intracellular ROS which can be counteracted by antioxidants such as N-acetylcysteine (NAC)[238]. NAC is an N-acetyl modified version of the amino acid cysteine that is thought to act primarily through increasing intracellular glutathione levels[239]. It has also been shown that NAC may directly reduce proteins by thiol-disulfide exchange[240,241]. It is possible therefore that, to some extent, the DTT treatment balances the increase in intracellular ROS that occurs in response to PDGF such that the redox balance of the cell is less perturbed than when treated with DTT alone. This would provide a mechanism by which to explain the apparently lesser response of cells to DTT when also treated with PDGF. It is possible that PDGF ultimately provides some protection against the effects of reductive stress caused by DTT alone.

However, it was noted that there was more variance in the quantification of proteins from PDGF treated samples than those treated with DTT alone. As discussed in section 5.2.3, this may mean that proteins that are changing in response to DTT treatment in the presence of PDGF are not deemed significant due to increased levels of variance. It would therefore be necessary to validate the differences between PDGF and DTT treated and DTT only treated samples to determine if the changes are indeed less with PDGF. It would also be interesting to investigate the potential protective effects of PDGF on the downstream cell survival to establish whether there is indeed a benefit to PDGF treatment at the same time as induction of reductive stress by DTT.

The findings that DTT treatment reduces the amount of ECM proteins, particularly collagen in fibroblast cells has potential relevance to clinical presentations. Fibrosis is a prevalent symptom of age related diseases, as well as a major debilitating effect of SSc, associated with erroneous deposition of excess collagen into the ECM. Fibrosis-inducing events and signalling cascades are known to generate intracellular ROS and oxidative stress[242]. The “western diet”, increased environmental toxicity and tobacco have all been implicated in increased pressure on our innate antioxidant system and it has been hypothesised than diet supplementation of antioxidants may provide some benefit[242]. A 2015 study by Conte et al., suggested some benefit of the anti-oxidant resveratrol in treating fibrotic diseases of the lung, although large scale clinical trials are yet to be

performed[243]. However, trials of antioxidants such as β-carotene, vitamin A and vitamin E as nutritional supplements have limited evidence of any beneficial effects[244]. Although nutritional supplementation may be of limited value for the treatment of fibrotic disorders, the potential use of reductants in topical treatments of dermal fibrosis to try and reduce erroneous collagen production should not be excluded.

These suggestions are built on the assumption that the reduction in ECM proteins is specific to reductive stress, and not also caused by oxidative stress. To determine whether the proteomic effects described in this chapter are indeed specific to reductive stress, it is necessary to make a comparison to oxidative stress. UV-A radiation makes up most of the ultraviolet light that comes from the sun and is therefore a major source of oxidative stress to the skin. The final results chapter of this thesis will therefore investigate the proteomic response to UV-A radiation-induced oxidative stress and compare these responses to those seen with DTT. The relevance of UV-A radiation to conditions such as skin cancer has led to extensive characterisation of the response of skin cells to this treatment. Thus, this investigation also provides opportunity to compare the proteomic responses seen in these quantitative proteomic analyses using the relatively new SWATH acquisition technique, to established responses already published in the literature. This will provide another level to assess confidence in protein identifications that can be obtained from variance analysis.

Chapter 6: The effect of ultraviolet-A (UV-A) radiation on the proteome of human dermal fibroblasts

6 The effect of ultraviolet-A (UV-A) radiation on the proteome of human dermal fibroblasts

6.1 Introduction

UV radiation emitted by the sun is categorised into three types according to wavelength. UV-A, which makes up 95% of the UV radiation that reaches the earth from sunlight, is classed as UV radiation with wavelength in the range 320-400 nm. Excessive exposure of the skin to harmful UV radiation is closely linked to the occurrence of skin cancers, with malignant melanoma accounting for 4% of all new cancer cases in the UK[245]. UV-B radiation, which causes direct damage to the DNA, has previously been assumed to be the main factor in skin cancers caused by sun-damage. However, in the last 15 years epidemiological evidence has arisen to suggest that in fact, the effects of UV-A radiation (the intensity of which varies with geographical location) are more closely linked to higher mortality rates[246]. As such, much research has been done to map specific effects of UV radiation on skin cells, particularly at the genetic level, to try and identify potential therapeutic targets. Whilst proteomic investigations have been done, there are no unbiased studies published that elucidate global proteomic responses to UV-A radiation, using quantitative gel-free technology.

UV-A radiation is absorbed by intracellular proteins or lipids including porphyrins and flavins for example riboflavin/FAD and haem, which then react to generate free radicals or singlet oxygen[123,247,248] . Increased ROS within cells leads to oxidative stress that has been associated with skin disease and photoaged phenotypes which include disrupted ECM, inflammation and increased apoptosis[249,250]. UV-A radiation has also been shown to stimulate an ER stress response in HDFs through XBP1 splicing and PERK mediated phosphorylation of eIF2α[122].

In contrast to the decreased ERK1/2 phosphorylation described in this thesis in response to DTT-induced reductive stress (Chapter 5), in some cell lines UV radiation has been shown to activate this signalling pathway as a way to promote survival and evade apoptosis[251]. UV radiation has also been shown to activate NFκB to promote release of pro-inflammatory cytokines, as well as stimulate expression of MMPs that lead to ECM degredation[252]. NFκB is a transcription factor that regulates the transcription of more than 60 pro-inflammatory genes involved in a range of functions including immune stimulation, apoptosis, cell adhesion and extracellular matrix degradation[253]. Under normal conditions NFκB is held in the cytoplasm by IκB proteins, which mask its nuclear localisation signal. Following UV-A radiation, these proteins are phosphorylated and targeted for degradation via the proteasome, allowing NFκB to transit to the nucleus to function[254,255]. The proteolytic degradation of ECM components, particularly collagens and elastic fibres, which is

associated with inflammation is one of the main recognised hallmarks of photodamaged skin. This leads to the typical wrinkled appearance of skin with reduced ability to resist stretching[256].

In this chapter, SWATH acquisition proteomics will be used to assess the global response of the BJ fibroblast proteome to UV-A radiation. Proteins that demonstrate a significant change will be tested for overrepresentation and enrichment of GO terms to assess whether any functional groups are disproportionately regulated. Responses of proteins associated with the extracellular matrix, and the breakdown thereof will also be investigated, and the global proteomic response compared to that described in chapter 5 for reductive stress.

6.2 UV-A radiation treatment for proteomic analysis

BJ fibroblasts were grown to 75% confluency before conditioned media was removed and cells treated with 5 J cm^{-2} UV-A radiation. Following treatment, the conditioned media was replaced and the cells left to recover for 18 h in a 37 °C humidified incubator. Conditioned media from control samples was also removed and dishes wrapped in tin foil for mock "exposure" to control for heating effects before recovery as with UV treated samples. Three biological replicates were set up, and the lysates used for three technical replicates of proteomic analysis by SWATH acquisition (Figure 6.1).

Following proteomic analysis, 3258 proteins were successfully quantified across the 9 total replicates. Of these, 515 proteins had a fold change of greater than 2 and FDR adjusted p-value of less than 0.05 when comparing control and UV treated quantifications. An increase, i.e. fold change of more than 2, was seen in 101 of the proteins whilst the remaining 414 proteins decreased with a fold change of less than 0.5.

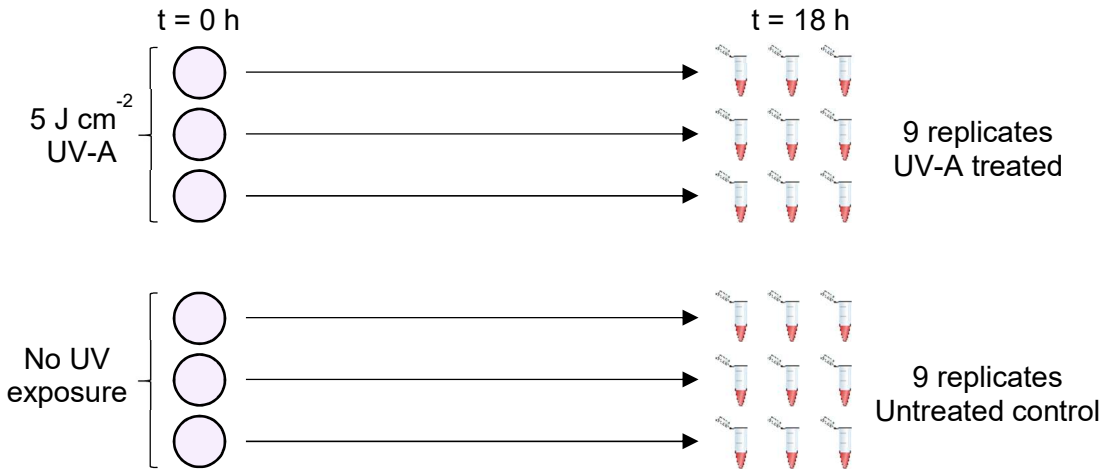


Figure 6.1: Outline of treatment of BJ fibroblasts with UV-A radiation. BJ fibroblasts were grown to 75% confluency before being subjected to 5 J cm^{-2} UV-A radiation. Prior to radiation conditioned media was removed and cells washed once in PBS. Control samples were then wrapped in foil to prevent exposure. Following treatment, conditioned media was replaced, and cells left to recover for 18 hours in a 37 °C humidified incubator before lysis in RIPA buffer containing phosphatase and protease inhibitors and analysis by quantitative LC-MS/MS (SWATH acquisition). Three biological replicates were set up, and three technical replicates were analysed from each of these leading to a total of 9 replicate samples.

6.3 Analysis of variance among replicates

The variation in quantification of proteins was analysed using the coefficient of variation (CV) and the standard deviation of log peak areas, as described for DTT and PDGF treatments in chapter 5 (see section 5.2.3). Whilst for the previous data 80 – 90% quantified proteins had variation within the desired range of these two measures, in this UV dataset only 50% of protein identifications met these criteria (Figure 6.2). There was, however, little difference in the spread of variation measured in UV-A or control samples, as was seen between DTT treated and PDGF and DTT treated samples in the previous chapter.

The additional variation could be due to several factors. With large numbers of proteins changing it may be that the cells response to UV-A radiation is more volatile and therefore there is a higher level of variation between biological samples. Alternatively, there may have been technical factors that impacted on the reproducibility of results. For example, where the ion intensity for a protein is close to the level of background noise, the FDR score from identification by comparison to the spectral library may be stochastically above or below the threshold for identification. Although SWATH acquisition is generally regarded as more reproducible than DDA, some variability between samples is nevertheless expected[257].

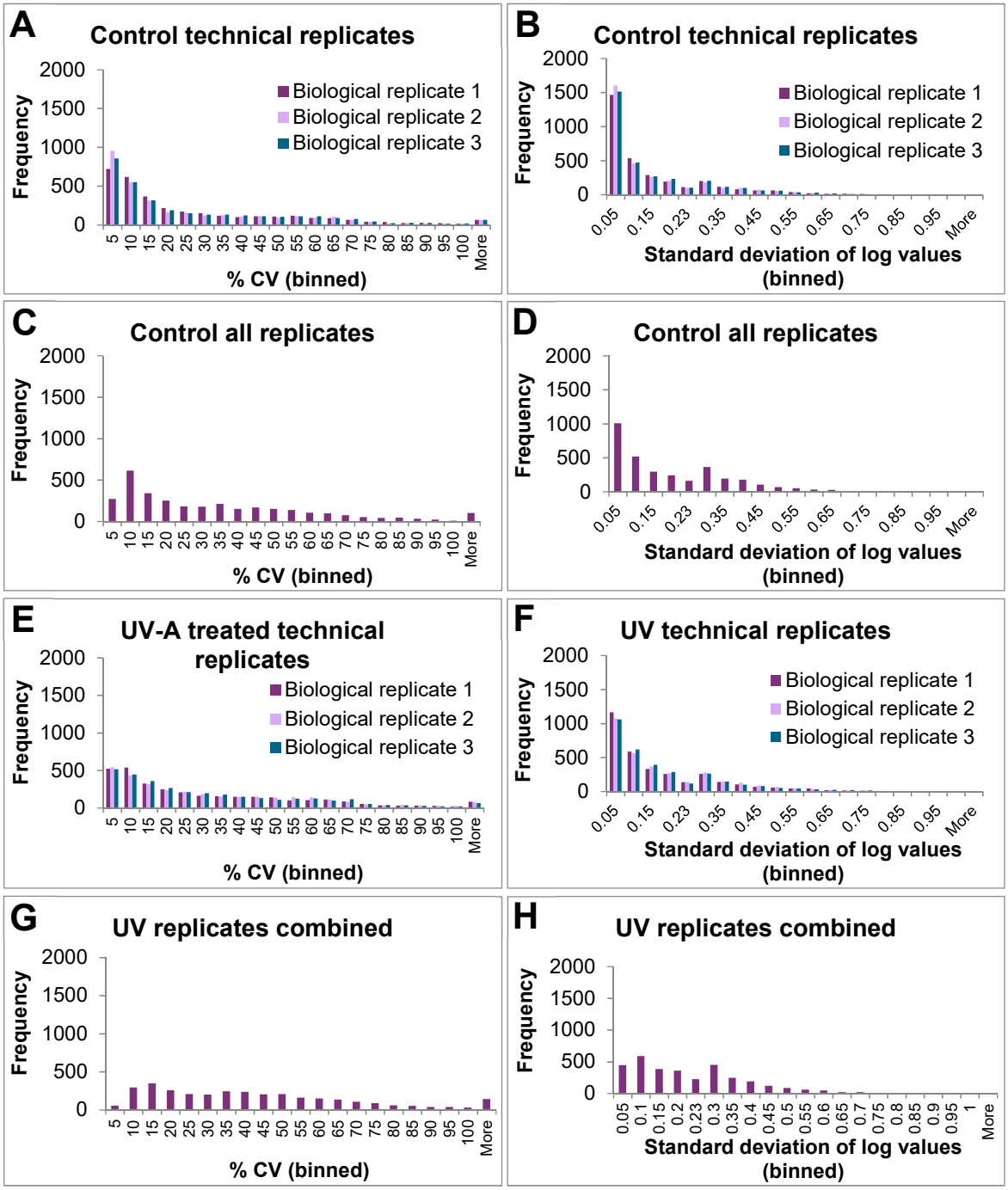


Figure 6.2: Histograms display distribution of coefficient of variance and standard deviation of log (peak areas) values across replicates. A – D show distributions of values in control samples and **E-H** show distribution of values in UV treated samples. **A** and **E** show %CV values from technical replicates; **B** and **F** show standard deviation of log values from technical replicates; **C** and **G** show %CV values from all replicates combined and **D** and **H** show standard deviation of log values from all replicates combined.

To try and identify whether there was a specific replicate or replicates that were different from the others a principal components analysis was performed (Figure 6.3). Principal components analysis (PCA) provides a way to reduce data with a high number of variables, such as that obtained from mass spectrometry analysis, into a series of principal components that demonstrate the biggest separations between samples. The plot of principal components 1 and 2 (PC1 and PC2) shown in Figure 6.3 re-plots the samples on two vectors (PC1 and PC2), across which samples show the biggest separation. As such the PCA provides a method by which it is possible to determine whether there are components of the data that can separate samples in to distinct groups as would be expected with a specific response to treatment.

The two conditions were separated well, with UV treated samples on the opposite side of the plot to the control samples. However, as expected, there was some obvious variation. The third biological replicate of UV-A treated samples (UV3) appeared somewhat separated from the main. The control samples were perhaps more closely clustered together, although a single technical replicate from biological replicate 1 (Cx1) sat away from the other technical replicates of that sample. It was therefore decided that there was not an obvious candidate or two that could be removed from the analysis without substantially reducing the data set. Instead, protein identifications were also filtered on the percentage CV value and those with %CV<30 across all replicates, both biological and technical, analysed separately from those that only met the fold change and p-value filters set previously.

The loadings plot for the PCA provides an indication as to which proteins are contributing the most to the principal component score. Loading values range between -1 and 1, with a value close to either extreme suggesting a substantial contribution of that variable on the corresponding principal component. Although there were no variables with a particularly high loadings score (i.e. >0.3), it was noted that a group of chaperone proteins including PDI (P4HB), as well as PPIB and SERPINH1 (two chaperone proteins heavily involved in the correct folding of collagen), lay outside the main cluster of proteins in the bottom right quadrant of this plot. This suggests that these proteins may be particularly involved in the separation of UV-A treated samples from control samples on the PCA plot. The two proteins with the highest loadings were serum albumin (ALB; top right quadrant) and vimentin (VIM; bottom right quadrant). The peak area for VIM in the UV-A treated samples was nearly double that in the control samples, indicating levels of vimentin are increased in response to UV-A treatment. Vimentin has previously been shown to play a role in the binding and stabilisation of collagen mRNAs[258] suggesting that, in contrast to the effects seen with DTT treatment, UV-A may lead to the stabilisation of collagen levels within these cells. This will be investigated further in section 6.5.2.

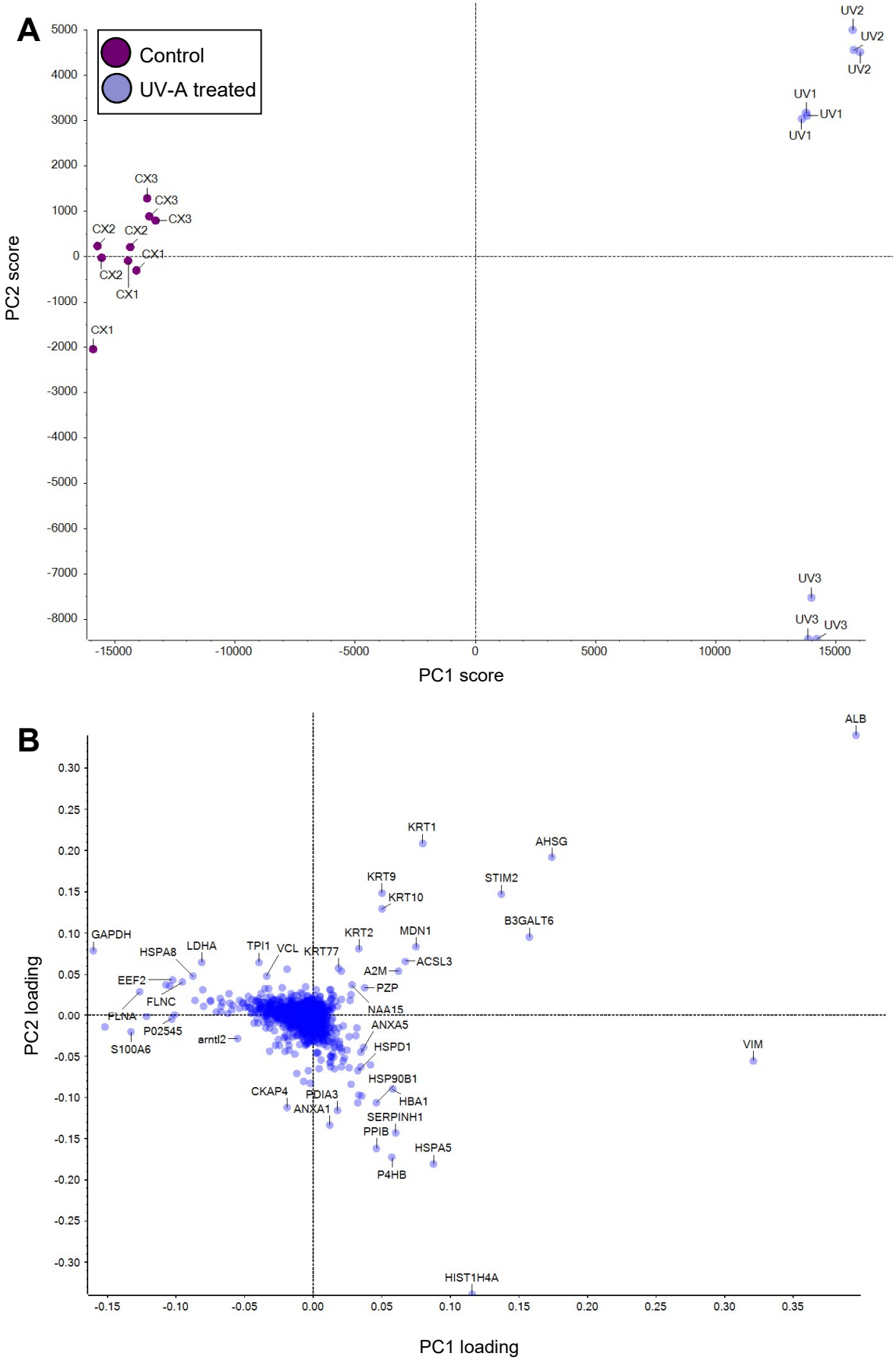


Figure 6.3: Principal components analysis of UV and control treated BJ fibroblast samples. A. Scores for PC1 (75.7%) versus PC2 (6.9%) with pareto scaling. Samples are labelled according to their biological replicate. **B.** Loading plot for PCA with outlying proteins labelled with gene name.

6.4 Proteomic responses to UV-A radiation

Proteins were plotted on a volcano plot to visualise the distribution of significant identifications following filtering for CV, fold change and FDR adjusted p-value (Figure 6.4). The number of proteins that met all three filtering criteria was 161, of which only 19 were increased compared to the control. This suggests the fibroblast response to UV treatment involved the downregulation of many more proteins than are upregulated. Even considering all proteins visible on the plot, there is an obvious bias towards the left-hand side, indicative of a general trend of decreasing relative protein amounts that were quantified in UV treated samples compared to the control.

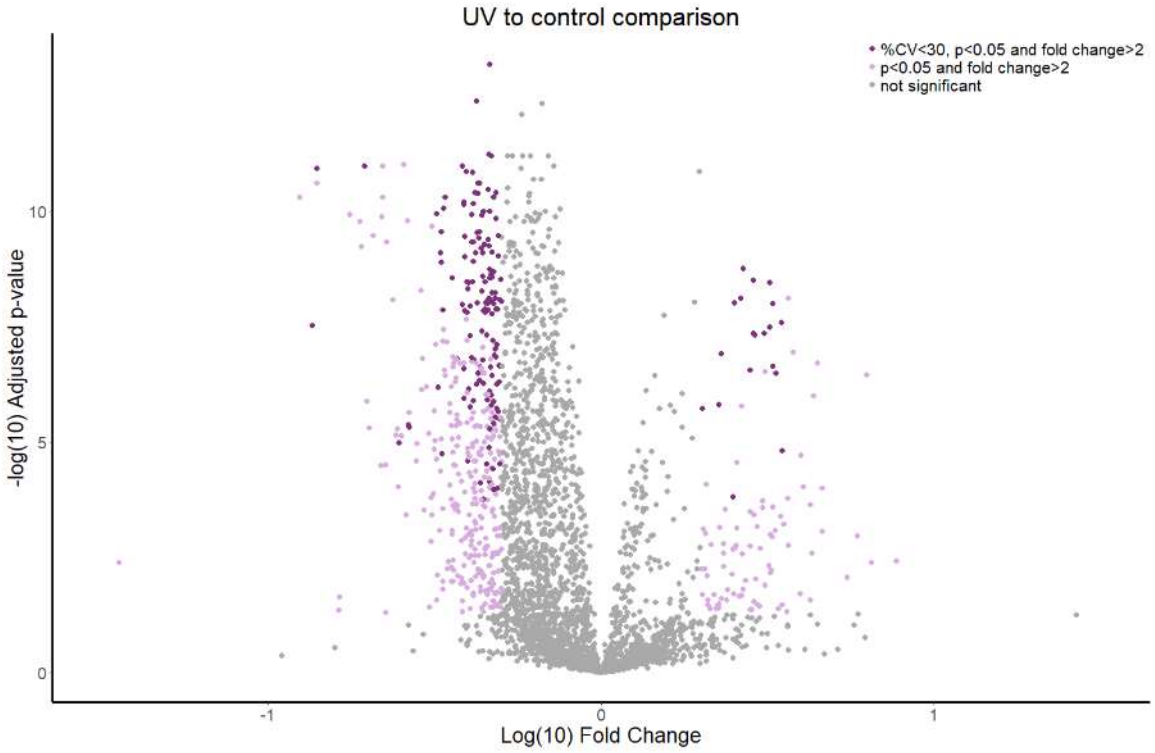


Figure 6.4: Volcano plot displays distribution of significantly changing proteins across fold change and significance values. Log_{10}Fold change is plotted against the negative log_{10}FDR-adjusted p-value for proteins quantified in the comparison of UV treated to control samples.

Copine-1 (CPNE1), an evolutionarily conserved calcium-dependent lipid-binding protein, was decreased in UV treated samples (fold change: 0.49; p-value: 3.09×10^{-6}). Ramsey et al., have previously showed that CPNE1 directs N-terminal proteolytic processing of the p65 subunit of NFκB[259]. NFκB is a transcription factor whose activity is known to be induced by UV radiation[252]. As described in the introduction to this chapter, its regulation is primarily through the cytosolic sequestering of NFκB by IκB proteins. The processing by CPNE1 renders NFκB unable to bind to DNA and therefore inhibits NFκB promoted transcription[259]. The decrease in relative CPNE1 amounts in the UV treated samples may suggest reduced NFκB inhibition, and therefore increased NFκB activity that is thought to occur in response to UV. NFκB target genes include cytokines and chemokines associated with the inflammatory response, as well as anti-apoptotic factors and cell cycle regulators that promote cell survival and proliferation[260]. DNAJA1 showed a decrease in response to UV-A treatment (fold change: 0.33; p-value: 2.31×10^{-9}). DNAJA1, whilst not known to be an NFκB target gene, has been associated with the promotion of apoptosis under stress conditions by inhibiting the JNK-mediated hyperphosphorylation of c-Jun, thus relieving the anti-apoptotic state. Its decrease, therefore, may serve as a protective and anti-apoptotic mechanism that could act synergistically with the expected survival responses of NFκB signalling.

Another transcription factor thought to be activated following UV radiation is p53. This is a tumour suppressor, and functions to promote transcription of a range of target genes including those associated with cell cycle arrest, DNA repair, and apoptosis[261]. Interrogating the data set presented in this chapter for the presence of p53 target genes resulted in the identification of 7 targets (Table 6.1). Unexpectedly, the target genes identified showed either no response (TRAF4), or a slight reduction in relative protein amount following UV radiation (all others). Latonen et al., have previously demonstrated a dose dependency of the p53 response to UV radiation, where low dose UV resulted only in a transient response by p53 and a high dosage was required for a sustained response[262]. In their study, Latonen et al., used 10 J cm^{-2} UV-B and UV-C radiation for their low dose, as opposed to the 5 J cm^{-2} UV-A that was used in the experiments presented here. Although UV-A is considered a more damaging wavelength due to its ability to penetrate further in to the skin, it is possible that this dose is still not sufficient to induce sustained p53 activation. Instead, the decrease seen in the relative amounts of these proteins may be anti-apoptotic, promoting cell survival rather than clearance of terminally damaged cells. The proteins proliferating cell nuclear antigen (PCNA), and ribonucleoside-diphosphate reductase subunit M2 B (RRM2B) are both involved in DNA damage repair following p53 activation. In the context of skin, it is possible then to consider the possibility that low dose UV-A exposure may in fact permit cells to maintain damaged DNA, with the potential to accumulate mutations that may eventually lead to carcinogenesis. In future experiments it would be interesting to compare the proteomic responses to different UV-A doses to fully investigate this possibility.

Table 6.1: p53 target genes. The UV versus control proteomic data set was interrogated for the presence of p53 target genes. Identified proteins are listed, along with their associated fold change and FDR adjusted p-value in the UV to untreated comparison. Statistically significant identifications (i.e. with an adjusted p-value < 0.05) are highlighted. Student's t-test, n=3.

Gene name	Function	Fold change	FDR adjusted p-value
TRAF4	Apoptosis	1.00	0.99
BAX	Apoptosis	0.84	7.89×10^{-4}
FDXR	Metabolism	0.84	0.53
TIGAR	Metabolism	0.41	1.27×10^{-5}
DDB2	DNA repair	0.76	0.53
PCNA	DNA repair	0.74	7.11×10^{-4}
RRM2B	DNA repair	0.53	0.011

Contrary to what might have been expected, there is limited evidence of an ER stress response amongst the proteins identified as increased in response to UV-A radiation. However, Stromal interaction molecule 2 (STIM2) was found to be higher in UV treated cells compared to control (fold change: 2.79; p-value: 2.82×10^{-7}; average %CV: 7.95). STIM2 is known to act as a sensor for ER Ca^{2+} levels, activating Ca^{2+} influx following depletion by re-locating to ER-plasma membrane junction sites and regulating plasma membrane Ca^{2+} channels to trigger Ca^{2+} influx. Farrukh et al., found that the ROS generated upon UV irradiation of fibroblast cells led to a rapid depletion of ER Ca^{2+} levels and induction of ER stress[263]. It is possible that the increase in the relative amount of this protein, which is seen in the data presented here, indicates a mechanism of cell recovery from oxidative stress induced by UV-A radiation. The recovery time allocated to the fibroblasts before lysis was 18 hours, which may have allowed for an ER stress response to be successful and regain homeostasis thereby reducing the potential for identifying proteomic changes more commonly associated with these pathways such as an increase in chaperone proteins.

Several cytoskeletal proteins were also identified as being significantly decreased following UV radiation (Table 6.2). The cytoskeleton primarily consists of microtubules, made up of tubulin proteins; and microfilaments, made up of actin proteins, as well as various other associated proteins involved in forming bridges between structures and providing transport along them. Previous studies have implicated UV radiation in the disruption of the cytoskeleton of fibroblasts, particularly with reference to microtubules[264,265]. These studies have suggested that UV-A radiation led to the destruction of microtubules and microfilaments, however the decrease in relative protein amounts presented in this thesis may suggest that this structural damage is accompanied by degradation of the proteins involved. Whilst tubulin proteins are amongst those proteins identified as significantly decreased, actin proteins themselves are not. Instead, it is mostly actin related or

filament associated proteins that are decreased in response to UV. The actin protein ACTA1 was found in the data set but with a fold change of 1.09 and adjusted p-value of 0.84. This might suggest that there are differing mechanisms behind the destruction of the different components of the cytoskeleton seen in response to UV, which may not have been fully investigated previously.

Table 6.2: Cytoskeleton associated proteins display a decrease in relative protein amount in response to UV radiation. Cytoskeletal or cytoskeleton-associated proteins are listed with their function, fold change and FDR-adjusted p-value in the UV to untreated comparison. Student's t-test, n=3. Data ranked by p-value.

Gene Name	Function	Fold change	p-value
MAP1A	Cross-linking between microtubules and other cytoskeletal elements	0.46	7.44E-05
TUBB3	Tubulin β3 chain	0.48	3.89E-06
TRIOBP	Actin cytoskeleton organisation	0.47	1.59E-06
PALLD	Actin cytoskeleton organisation	0.32	6.53E-07
CFL1	Actin filament depolymerisation	0.48	1.41E-07
PFN2	Actin binding	0.45	4.70E-08
TUBB6	Tubulin β6 chain	0.47	7.61E-10
TUBB	Tubulin β chain	0.47	7.61E-10
TUBB4B	Tubulin β4B chain	0.42	2.81E-10
PDLIM1	Protein adaptor; stress fibre formation	0.44	1.01E-10
LASP1	Regulation of actin cytoskeletal activity	0.48	3.91E-11
CNN2	Filament associated protein	0.42	3.91E-11
DSTN	Actin filament depolymerisation	0.47	6.34E-12
DYNC1LI2	Accessory protein for vesicle transport along microtubules	0.46	5.79E-12
CNN3	Filament associated protein	0.42	4.09E-13

6.4.1 The STRING-db tool identifies protein-protein interactions with groups of proteins that show a significant response to UV-A treatment

The STRING-db tool was used to determine if any of the proteins that showed a significant response to UV interacted with each other (Figure 6.5 and Figure 6.6). Proteins that displayed a fold change greater than 2, with an adjusted p-value less than 0.05, were entered as input for multiple protein analysis in STRING version 10.5 and tested for experimentally determined interactions, or those listed in curated databases. Disconnected nodes, which were those proteins who did not interact with any others in the input, were not displayed.

Amongst the proteins that showed an increase in response to UV there were only a few, usually single stranded, connections between a subset of the proteins (Figure 6.5). These proteins were all associated with serine proteinase or endopeptidase inhibition. Serine proteinases are distinct from the MMPs that are classically thought of as ECM degradative enzymes. The serpin family of serine proteinase inhibitors, of which two members are implicated in the network depicted in Figure 6.5, have been shown to inhibit the activity of a recently described membrane-anchored serine protease matriptase-3[266]. Matriptase-3 has been shown to promote tumour growth where a cancer cell line transfected with matripase-3 produced subcutaneous tumours in mice of a higher mass than those formed from mock transfected cell implants[267]. An increase in the inhibition of this, and other serine proteases, may therefore suggest protection against excessive proteinase activity. Further evidence for these identifications being associated with a protective response is gained from the GO annotations associated with these proteins. ALB is also annotated to the GO term ‘negative regulation of apoptotic process’ and ‘negative regulation of programmed cell death’, and SERPING2 is annotated to ‘positive regulation of collagen biosynthesis’. This suggests a role for these proteins in the recovery of damaged ECM and protection against apoptosis that may otherwise occur from lack of cell-matrix contacts.

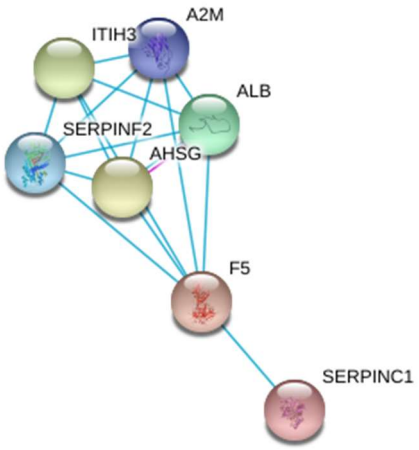


Figure 6.5: Protein-protein interactions amongst proteins increased in response to UV determined using the STRING-db tool. The multiple protein search function of the STRING-db tool found at string-db.org was used to identify interactions between proteins identified as significantly decreased in response to UV radiation. This identified a small cluster of interacting proteins involved in serine proteinase inhibition. Known interactions from curated databases (**turquoise**) or experimentally determined (**pink**) with the highest confidence (minimum required interaction score 0.9) were used. Disconnected nodes were hidden. Produced using STRING version 10.5.

In the significantly decreased set of proteins, this analysis identified a cluster of interacting proteins (highlighted with a purple box in Figure 6.6) that were all associated with mRNA splicing, processing or maturation in some way. Synthesis of mRNA is repressed in response to DNA damage and therefore the lower abundance of these proteins is predictable. It is interesting to contrast the decrease in p53-mediated DNA damage repair proteins identified previously (Table 6.1) with the downregulation of mRNA synthesis-associated proteins that is often associated with DNA damage. It might be expected that DNA damage repair would be increased while mRNA synthesis is repressed. However, the depletion of hnRNP H/F, two of the proteins identified in this data set, has also been shown to compromise p53-mediated apoptosis[268]. hnRNPs are known to function in the processing of heterogeneous nuclear RNA to mature RNA, and stabilising this for transport. They are also involved in the regulation of mRNA alternative splicing events[269]. hnRNP H1 has been implicated in regulating the splicing of MAPK and ubiquitin pathway genes[270], and its loss may be associated with a more oncogenic phenotype in breast cancer cells[271]. It is possible, therefore, that these responses are focussed on promoting cell survival rather than damage repair. Different aspects of the UV response are likely to occur at different time points after exposure, and it would be interesting to try and isolate these stages of recovery in future work.

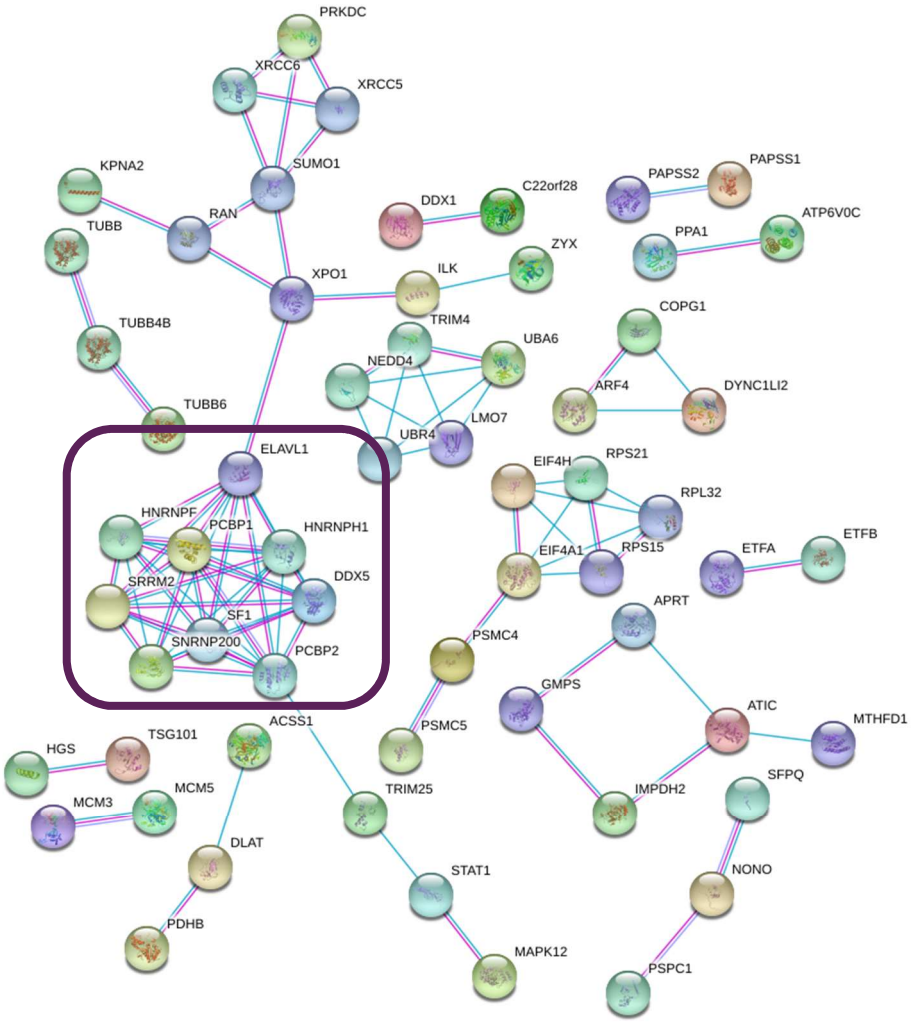


Figure 6.6: Protein-protein interactions amongst proteins decreased in response to UV determined using the STRING-db tool. The multiple protein search function of the STRING-db tool found at string-db.org was used to identify interactions between proteins identified as significantly decreased in response to UV radiation. This identified a cluster of interacting proteins involved in mRNA splicing, transport or transcription (highlighted by purple box). Known interactions from curated databases (**turquoise**) or experimentally determined (**pink**) with the highest confidence (minimum required interaction score 0.9) were used. Disconnected nodes were hidden. Produced using STRING version 10.5.

Overall, the responses seen here in response to UV radiation are indicative of a protective response against the effects of UV-A induced damage to the cell layer. Cell survival, rather than apoptosis, seems to be promoted and excessive ECM degradation protected against by increased levels of proteinase inhibitors. In the following section, identifications are tested for significant overrepresentation or enrichment of specific GO terms to gain further insight in to the predominant responses seen in this data.

6.4.2 Statistical overrepresentation is seen predominantly in proteins upregulated in response to UV-A radiation

Testing the increasing and decreasing proteins for statistically significant overrepresentation or enrichment of GO terms revealed surprising results (Figure 6.7). Although most proteins seemed to decrease in response to UV treatment, overrepresentation was much more prevalent in the increasing proteins. Significant GO overrepresentation was only seen when testing the decreasing proteins against the PANTHER GO Slim database, without filtering for %CV. Under these conditions 'mRNA splicing, via spliceosome' displayed overrepresentation, compared to all proteins identified in the data set, of 257%. This means that there were 2.5 times as many proteins annotated to this term as would be expected based on the distribution of terms within the whole data set. This overrepresentation accords with the cluster of proteins associated with mRNA splicing seen in and discussed in the analysis of decreasing proteins using the STRING database above (Figure 6.6).

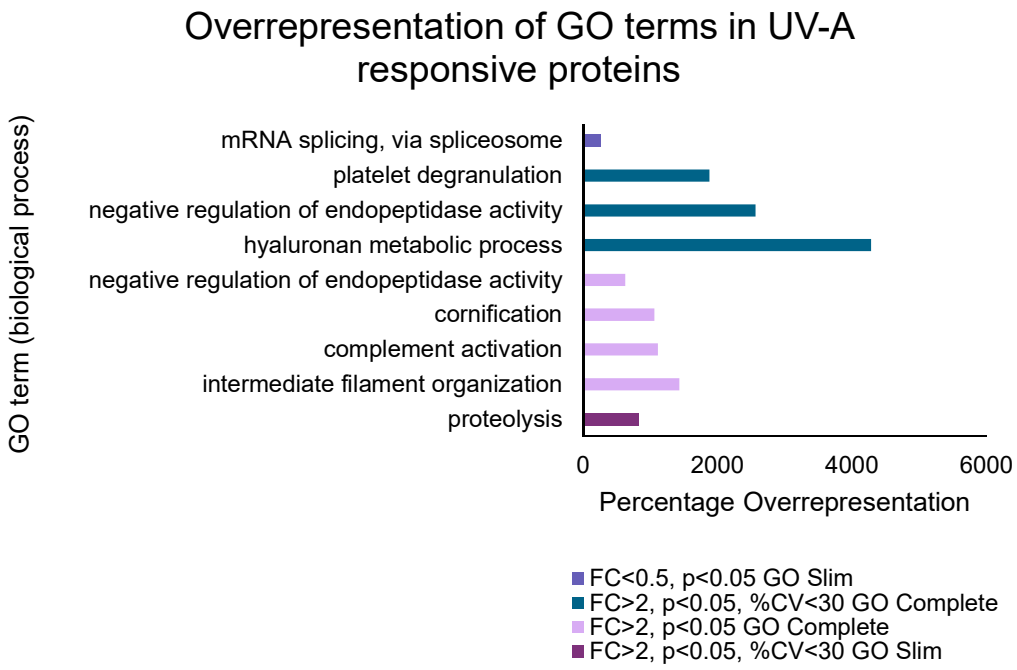


Figure 6.7: Analysis of significantly changing proteins in response to DTT reveals statistically significant overrepresentation of GO terms. Percentage overrepresentation is plotted for GO biological process terms found to be statistically significantly overrepresented in those proteins that changed in response to UV treatment. Significantly changing proteins were tested against a reference list of all quantified proteins in the data set using PANTHER Overrepresentation Test (Released 20171205) and PANTHER version 13.1 Released 2018-02-03. A Fisher's exact test with FDR multiple test correction was carried out using the GO biological process complete or the PANTHER GO Slim annotation set. For ease of viewing, where both parent and daughter terms were found to be overrepresented, only the parent term is displayed.

Overrepresentation of GO terms was seen in more cases when looking at the proteins that increased in response to UV-A radiation. With no filtering for %CV, overrepresentation was seen in the GO complete database for 'negative regulation of endopeptidase activity', 'cornification', 'complement activation' and 'intermediate filament organisation'. With %CV filtering, overrepresentation was seen with the GO Slim database for 'proteolysis' and with the GO complete database for 'platelet degranulation', 'negative regulation of endopeptidase activity', and 'hyaluronan metabolic process'.

'Negative regulation of endopeptidase activity', which was overrepresented in both %CV filtered and unfiltered results using the GO complete database, was annotated to the same proteins that mapped to 'proteolysis' using the GO slim database. These proteins were a series of serine protease inhibitors including the inter-alpha-trypsin inhibitor family members H1, 2 and 3. These serine protease inhibitors are involved in the stabilisation of the extracellular matrix, in part by binding to hyaluronan[272], and can inhibit tumour progression and metastasis of cancerous cells[273]. Their upregulation suggests an ECM stabilisation response that may be protective against the direct damage to ECM proteins caused by UV-A radiation[274].

Also annotated to these terms, as well as to 'complement activation', was complement component 3 (C3). It has been suggested that the alternative complement pathway may act as a UV sensor, leading to the upregulation of C3 as has been seen here in this data (fold change: 2.62; p-value: 7.58×10^{-9}; average %CV: 12.65)[275]. Other components of the terminal complement complexes C5 and C9 were also mapped to 'complement activation' in the test without %CV filtering. CD59, an inhibitor of the complement membrane attack complex (MAC), shows a very slight but significant decrease in response to UV (fold change: 0.90; p-value: 0.034; average %CV: 1.53). These suggest the complement pathway may be upregulated in response to UV-A radiation, perhaps to clear apoptotic cells.

Functional enrichment testing was also carried out on the data set to determine whether the proteins annotated to any GO terms displayed a statistically significant difference in distribution of fold changes, compared to the distribution of all proteins identified. No enrichment existed for biological process terms; however, proteins annotated to the cellular compartment term "extracellular region" were enriched with a p-value of 0.017. Many of these proteins were either ECM components themselves, or protease inhibitors as discussed previously. The distribution of fold changes for these proteins was shifted towards higher values compared to the distribution of fold changes for all identifications (Figure 6.8) suggesting an increase of extracellular proteins in response to UV-A radiation treatment. This provides further evidence that, in cells treated under the experimental conditions described in this chapter, there is perhaps a protective or damage

limitation response whereby ECM components are being secreted to re-populate or support a potentially damaged ECM and protected from degradation by proteinase inhibitors.

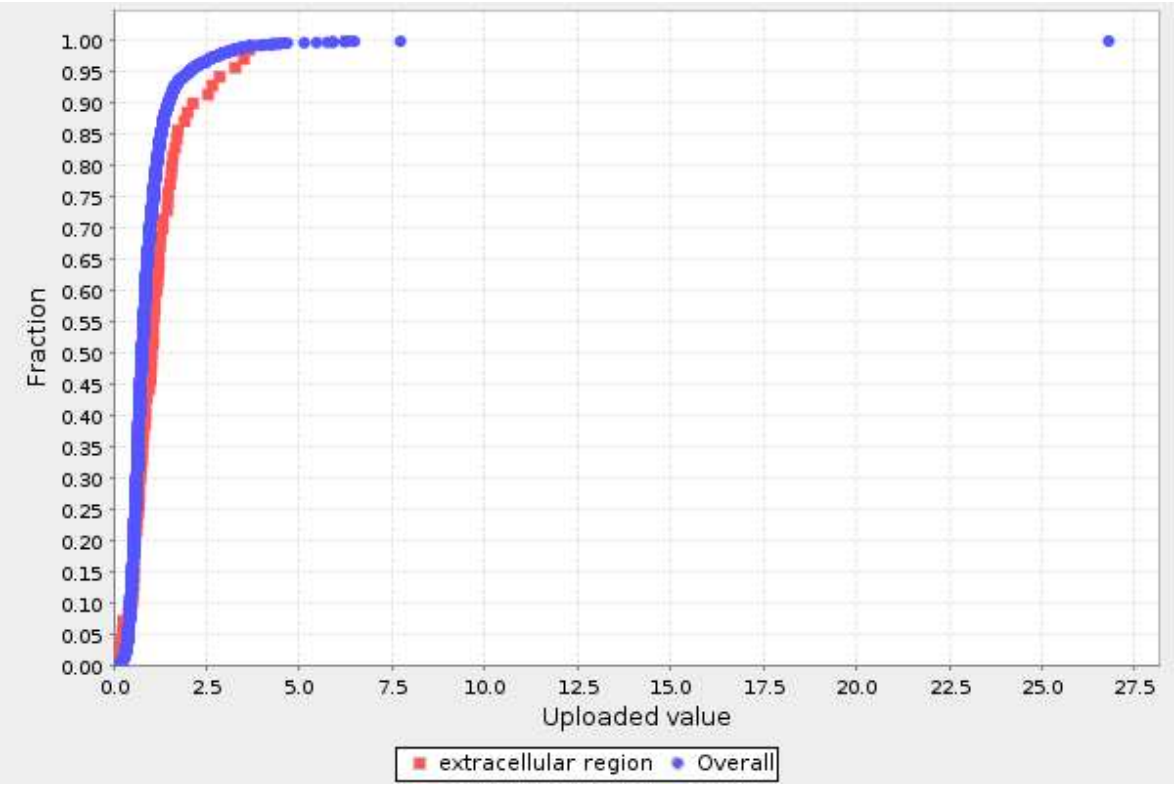


Figure 6.8: Extracellular region is significantly enriched in UV-A treated to Untreated comparison. Cumulative fraction plotted against uploaded value of fold change for all proteins quantified in UV-A treated to untreated comparison. Distribution of overall values shown in blue, and distribution of values annotated to extracellular region shown in red. PANTHER Enrichment Test (Released 20170413) was used with PANTHER version 13.1 (Released 2018-02-03) and PANTHER Go-Slim cellular compartment annotation data set. Bonferroni correction for multiple testing used for significance values. “Extracellular region” was found to be significantly enriched with a p-value of 0.017.

6.5 Comparison of UV-A and DTT responding proteins

Having surveyed the proteomic changes that occurred following UV-A radiation, these responses were then compared to the changes seen in DTT treated cells (Chapter 5). Both treatments will have caused stress due a redox imbalance in the cell: DTT acting as a reductive stress and UV-A as an oxidising stress. It was therefore interesting to determine whether the responses are common, and therefore likely to be general responses to redox stress, or specific to perturbations of a reducing or oxidising nature.

6.5.1 A small subset of proteins changes significantly, and similarly, in response to both DTT and UV-A radiation treatments

Of the 45 proteins that were significantly changed in response to DTT treatment (without PDGF), 9 were also significantly changed in response to UV-A radiation (Table 6.3). All these proteins were downregulated in both cases, though in some cases the downregulation was greater with UV treatment than DTT.

Table 6.3: Comparison of fold change and p-values for proteins identified as significantly changing in both DTT and UV treated data sets. The fold change, p-value and average %CV is displayed for the proteins identified in the DTT and UV treated data sets as significantly changing (ie. fold change > 2; p-value < 0.05).

Peak Name	Gene Name	DTT			UV		
		Fold change	p value	%CV	Fold change	p value	%CV
P29279	CTGF	0.41	5.50E-06	22.15	0.40	1.77E-03	40.99
Q9HCU0	CD248	0.49	1.03E-06	14.73	0.34	4.90E-11	10.88
Q8WUJ3	CEMIP	0.39	1.88E-08	14.88	0.30	6.65E-06	37.17
Q9BZL1	UBL5	0.49	2.22E-02	39.43	0.29	7.43E-07	39.48
Q6EMK4	VASN	0.50	3.59E-07	11.76	0.25	9.22E-05	39.54
P07686	HEXB	0.47	8.01E-04	19.80	0.22	4.90E-11	24.30
O00622	CYR61	0.26	5.10E-05	25.02	0.22	1.02E-11	19.49
P48729	CSNK1A1	0.48	2.62E-02	41.09	0.20	1.33E-06	65.68
P35625	TIMP3	0.21	9.25E-07	24.55	0.14	2.46E-11	28.26

TIMP3 was downregulated in both DTT and UV treated cells, although the downregulation was more pronounced in UV treated cells (fold change: 0.14; p-value: 2.46×10^{-11}), equivalent to a 7-fold decrease in relative protein amount. TIMP3 irreversibly inactivates MMPs and therefore its decrease is suggestive of an environment that supports increased MMP activity, such as that expected to occur in response to UV radiation. The regulation of this protein indicates that both treatments support increased MMP activity. This contrasts with the increase in serine proteinase inhibitors seen in UV-A treated cells and suggests that the distinct classes of proteinases that act on the ECM are regulated differently in response to treatment.

Several of the other proteins downregulated in response to both DTT and UV-A treatment are associated with the promotion of angiogenesis, cell proliferation and cell adhesion (CYR61, CTGF, and CD248) suggesting that these activities may be inhibited. CD248, otherwise known as endosialin, has been shown to bind to fibronectin (FN1) as well as type I and type IV collagen proteins[276]. One study has shown that cells overexpressing CD248 adhere faster to ECM-coated culture surfaces than controls[276]. As there was only a moderate decrease in the relative amount of fibronectin that was seen in response to either treatment (UV fold change: 0.85, p-value: 0.02; DTT fold change: 0.63, p-value 0.0043), it is possible that reduced cell adhesion seen in response to treatment is partly mediated by reduction in proteins such as CD248.

A study published in 2000 demonstrated that CYR61 and CTGF mediate cell adhesion through binding to integrins, including integrin $\alpha_6\beta_1$ and cell surface heparin sulfate proteoglycans (HSPGs)[277,278]. More recently, fibroblast binding of integrin $\alpha_6\beta_1$ and the HSPG syndecan-4 to CYR61 has been shown to induce apoptosis, in contrast to the enhanced cell survival displayed when binding to endothelial cells[279]. It is possible, then, that the substantial decrease in the relative amount of this protein quantified in response to either DTT or UV treatment acts as a protective mechanism to reduce apoptosis across the cell population.

Finally, CEMIP showed a fold change of 0.3 in response to UV radiation (indicative of a 3-fold decrease in relative protein amount). CEMIP mediates the depolymerisation of hyaluronan via the cell membrane-associated clathrin coated pit endocytic pathway[280]. The increase in inter-alpha-trypsin inhibitor family members H1, 2 and 3 following UV-A treatment was discussed in GO overrepresentation analysis and these findings combined suggest an environment of increased hyaluronan stabilisation. Hyaluronan usually has a rapid turnover but it is implicated in wound healing, promoting fibroblast migration to the wound site, and as a major component of ECM in wound closure[281]. It has also been implicated in aiding the migration of inflammatory response cells to wound sites[281]. It is possible, therefore, that stabilisation of hyaluronan in response to treatment are part of a wound healing response to allow recruitment of fibroblasts and other cells to the site of damage.

6.5.2 Collagen proteins respond differently with DTT and UV-A radiation treatments

Collagen proteins were identified in the DTT treatments as significantly decreased, though often just below the threshold of 2-fold change set during the initial proteomic analysis. To compare the response of collagen proteins in UV-A treated cells to that seen in DTT treated cells, the data set was interrogated for collagen proteins. Many collagen proteins were also identified in the UV treated samples, however the fold changes seen in response to UV-A treatment centred more around a value of 1, suggesting a more limited response than those seen in response to DTT (Figure 6.9). The only collagen protein that showed a fold change of greater than 1.5 in the UV treated sample was COL1A1 (Table 6.4). Interestingly, DTT treatment induced an almost opposite response: UV-A treated cells showed an increase of this protein compared to untreated (fold change: 1.52; p-value: 2.8×10^{-5}; %CV: 12.3) whereas DTT treated cells showed a decrease compared to untreated (fold change: 0.57; p-value: 2.25×10^{-3}; %CV: 19.48). This suggests that the decrease of collagen proteins is unique to the reductive stress response in BJ fibroblasts.

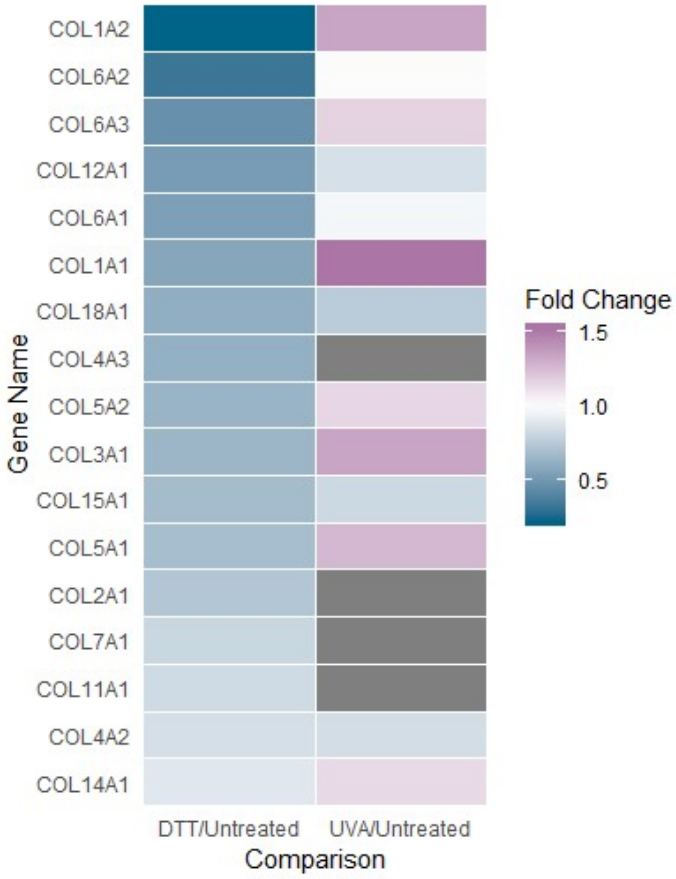


Figure 6.9: Fold changes of collagen proteins are greater with DTT treatment than with UV-A radiation. The fold change of collagen proteins identified in the DTT treated data set is plotted as a heatmap to compare the responses to DTT treatment with UV-A irradiation. Collagen proteins which were not identified in the UV-A treated data set are coloured grey. Data ranked by fold change in DTT/Untreated comparison.

Table 6.4: Comparison of fold change and p-values for collagen proteins identified in DTT and UV treated data sets. The fold change, p-value and average %CV is displayed for the collagen proteins identified in the DTT treated data set. These are compared to the values associated with the same proteins, where present, in the UV treated data set. Fold changes of greater than 1.5 in either direction are highlighted. Data ranked by gene name (alphabetical).

Gene Name	DTT			UV		
	Fold Change	P-value	%CV	Fold Change	P-value	%CV
COL11A1	0.83	0.85	58.40			
COL12A1	0.52	5.94E-10	5.55	0.85	4.86E-03	8.33
COL14A1	0.89	0.94	89.17	1.14	0.67	44.51
COL15A1	0.68	0.28	29.51	0.82	0.83	135.06
COL18A1	0.61	4.08E-04	15.04	0.76	9.05E-03	16.15
COL1A1	0.57	2.25E-03	19.48	1.52	2.80E-05	12.30
COL1A2	0.22	1.25E-06	20.08	1.33	1.53E-04	10.15
COL2A1	0.73	0.67	51.26			
COL3A1	0.65	0.02	18.33	1.33	2.95E-04	11.41
COL4A2	0.85	0.77	36.06	0.84	0.35	28.53
COL4A3	0.62	0.25	5.29			
COL5A1	0.69	2.24E-03	12.29	1.27	1.05E-03	9.24
COL5A2	0.64	5.23E-03	19.11	1.15	0.63	44.72
COL6A1	0.54	1.20E-03	18.87	0.96	0.44	7.95
COL6A2	0.33	9.38E-07	15.33	1.01	0.80	8.10
COL6A3	0.46	7.05E-08	6.96	1.16	1.13E-03	6.12
COL7A1	0.81	0.80	56.25			
COL8A1	1.20	0.69	30.88			

Photoaged skin is usually associated with a loss of collagen, however the relative amount of collagen proteins themselves in this data set show little change in response to UV at the recovery timepoints tested here. A decrease in the relative amounts of TIMP proteins was seen (see section 6.5.1) which may suggest an environment that is supportive of MMP collagenase activity in the ECM, although the MMPs associated with this activity were not detected. As discussed earlier in this chapter with reference to ER stress (see section 6.4), it is possible that 18 hours recovery after exposure permitted cells enough time to begin to repair any damage caused or was not sufficient to detect longer-term changes in more stable/long-lived proteins. A study focussed on the recovery of the microtubule cytoskeleton in mouse fibroblasts exposed to UV radiation found that after 48 hours, recovery was complete and cells showed normal morphology[265]. It is therefore possible that, at the timepoints considered in the study presented in this thesis, some recovery mechanisms to sub-lethal UV radiation are already in place.

6.6 Discussion and Conclusion

This chapter has described the proteomic response of BJ fibroblasts to UV-A radiation and demonstrated both similarities and differences between this and the response to DTT described in chapter 5. Whilst UV-A radiation appeared to alter many more proteins than DTT treatment (515 significantly changing proteins compared to 45 with DTT) there were fewer significantly overrepresented GO terms associated with these identifications. This perhaps indicates a more global protein response than one associated with the reaction of a specific set of proteins. This may be because the ROS generated in response to UV-A radiation affects many aspects of the cell physiology including damaging the DNA, oxidising membrane phospholipids as well as inducing a general oxidative and ER stress response.

ECM degradation has previously been implicated as a major response to UV-A radiation through upregulation of proteases that damage key proteins such as collagens and elastic fibres[282,283]. In this data set there was a distinction between the responses of serine proteinase inhibitors, which increased in response to UV-A radiation, and metalloproteinase inhibitors (TIMPs), which decreased. The two classes of proteinases on which these inhibitors act have different targets within the ECM suggesting a finer tuned response in the ECM than uniform degradation.

Stabilisation of the ECM protein hyaluronan appeared from several angles of analysis, which suggested that this may be a distinct response of these cells to UV-A radiation, not recognised before in the literature. Both a significant decrease in the hyaluronan depolymerising protein CEMIP, and an increase in the inter-alpha-trypsin inhibitor family members H1, 2 and 3 which are involved in ECM stabilisation suggest a response that promotes the stability of hyaluronan following UV-A radiation. This may help to counteract the degradation of other ECM components seen, as well as acting as a signal to promote inflammation and wound healing[281]. CEMIP was also decreased in response to DTT treatment, however no significant change was observed in the inter-alpha-trypsin inhibitor family proteins suggesting that the stabilisation of hyaluronan may be specific to the UV-A response.

The behaviour of collagen proteins after UV-A radiation was distinct from that seen with DTT treatment in the previous chapter. It was noted that the recovery times afforded to the cells under different treatment conditions were not the same and it is possible that, in the UV-A treated cells, some level of recovery has already occurred. In future studies it would be interesting to titrate both the dose- and time- dependency of the proteomic responses to UV-A radiation and further compare these to the reductive stress response. Similarly, it would be interesting to look at the secretome of UV-A irradiated fibroblasts to assess whether the ECM response is more clearly visible.

This data has provided evidence that the responses of BJ fibroblasts to DTT, including the relative decrease in collagen proteins, could be specific to reductive stress. As little research has been done on the specific effects of reductive stressors on dermal skin cells, it is interesting to consider the effects of reductants on skin physiology and the potential relevance of changes in collagen composition to the skin care market for both medical and cosmetic purposes. The final chapter of this thesis will discuss the results described in these chapters and their potential impact on the function of dermal fibroblasts in the skin. Suggestions of future study will also be made to further expand upon these findings.

Chapter 7: Discussion

7 Discussion

This thesis has described the responses of human dermal fibroblasts to DTT-induced reductive stress from both a signalling and proteomic perspective. The studies revealed a previously undescribed growth factor-independent activation of chronic Akt phosphorylation in response to DTT, which was accompanied by a dephosphorylation of the extracellular signal regulated kinases p44/42. This response appeared to be specific to reductive stress and was not seen following treatment with redox-independent ER stress reagents. In addition, DTT treatment induced a reduction in the relative collagen amounts detectable in cell lysates, which was subsequently demonstrated not to occur in response to oxidative stress induced by UV-A radiation.

This chapter will discuss the novel signalling responses to DTT, and the potential modulation of this response by PDGF. The effects of reductants on the ECM will then be discussed in contrast to the effects of UV-A radiation on the same. Finally, the use of SWATH MS as a proteomic tool will be discussed with reference to the data presented here, and suggestions made for future work. The conclusions drawn from the work presented in this thesis are summarised at the end.

7.1 DTT induces a growth factor independent signalling response

The treatment of cells with DTT led to a chronic phosphorylation of Akt, independently of growth factor stimulation, which has not been previously described in the literature (Figure 4.9). A study by Murata et al., has previously reported growth factor-independent, long-term Akt phosphorylation following treatment with hydrogen peroxide in cardiac myocytes transfected with glutaredoxin[284]. They went on to demonstrate the anti-apoptotic effect of this response and suggested it was due to the modulation of the redox state of Akt by glutaredoxin[284]. A disulfide bond between Cys297 and Cys311, in the activation loop of Akt, has been shown to be sensitive to reductants, being broken by incubation of the protein with 20 mM DTT[285]. The formation of this disulfide bond was shown to be accompanied by an increased association of protein phosphatase 2A (PP2A), which dephosphorylates Akt, preventing its long-term signalling[284]. Glutaredoxin prevented the formation of this disulfide bond and therefore reduced the association of PP2A allowing phosphorylation of Akt to persist[284]. It is possible that the stimulation of chronic Akt phosphorylation by DTT that was seen in this thesis may occur in a similar manner, with DTT reducing and preventing the formation of the Cys297-Cys311 disulfide bond.

The study by Murata et al., suggests that chronic Akt phosphorylation is associated with an anti-apoptotic response. As discussed in chapter 4, chronic Akt phosphorylation has also been associated with the establishment of senescence[206,207], a cellular phenotype where cells have resisted apoptosis to become stable but under a state of permanent growth arrest (See section 4.4.2). Although there was no conclusive evidence for induction of senescence in the studies presented in this thesis, it is possible that chronic Akt phosphorylation may be a signalling response that is beneficial to cell survival be that in a way that is senescence inducing, or simply anti-apoptotic.

7.1.1 PDGF modulates the response to DTT

There are distinct differences in the proteomic response of cells to DTT treatment depending on whether PDGF is present or not. This finding provides evidence of a modulatory effect of PDGF on the cells' response to reductive stress induced by DTT, though the precise nature of this is not yet clear.

Treatment of BJ fibroblasts resulted in chronic phosphorylation of Akt and dephosphorylation of p42/44, as observed over an extended time course. However, in the presence of PDGF, the Akt phosphorylation decreased towards the end of the time course, and the dephosphorylation of p42/44 occurred from an earlier point (Figure 4.12). As described above, chronic Akt phosphorylation has been associated with an anti-apoptotic response. However, these published studies did not stimulate the cells for as long as 6 hours and it is possible that after a certain time

point Akt phosphorylation is no longer beneficial. Inhibition of Akt with perifosine induced a small but insignificant increase in the proliferation of cells following DTT treatment (Figure 4.13), which may indicate that the reduced Akt phosphorylation, seen in later time points with PDGF and DTT together, benefits cell recovery and proliferation.

Alternatively, there is some evidence that suggests downregulation of both Akt and ERK may increase apoptosis. Jin et al., demonstrated that where apoptosis is a result of tumour necrosis factor-related apoptosis-inducing ligand (TRAIL), expression and activation of the apoptosis proteins Bax, Bid and caspase-3 was mediated by decreased ERK and Akt signalling[286]. Therefore, there is a possibility that the decrease in Akt signalling seen in the present study may be an effect of PDGF sensitising cells to the damaging effects of DTT and thus promoting apoptosis. Further insight in to this was gained from the proteomic data.

Statistically significant overrepresentation of the GO term 'regulation of ERK1 and ERK2 cascade' was seen among those proteins that demonstrated a significant change in response to DTT treatment (Figure 5.13). The proteins associated with this term were decreasing with DTT treatment, which would be linked to a decrease in ERK1/2 (p42/44) phosphorylation, in agreement with the signalling responses seen. There was little or no difference between the fold changes of these proteins seen in response to DTT alone, or PDGF and DTT together, suggesting a broad similarity in the response under the two conditions. However, there was a lesser decline of collagen proteins in response to DTT seen with PDGF and a greater, and significant, increase in TIMP1. These results suggest that the addition of PDGF lessens the degradation of collagen ECM in response to DTT, which is likely to be beneficial to the cells.

The significant decrease of mitochondrial iron-sulfur cluster assembly enzyme ISCU, seen after PDGF and DTT combined treatment, suggests there may be some responses occurring that reflect the combined effects of both treatments. This ISCU protein was significantly decreased in PDGF and DTT treated cells compared to DTT or PDGF only and decreased (though insignificantly) when compared to untreated cells, suggesting the response only occurred in response to both treatments together. A decrease in ISCU has been associated with a decrease in Krebs cycle enzyme activity and mitochondrial function, leading to an increase in ROS generation[287]. ROS generation from decreased ISCU has been associated with increased apoptosis of cells[288]; however it may also be acting to rebalance the redox state of cells in response to DTT mediated reductive stress.

Overall, whilst there is clear evidence of PDGF modulating the effect of DTT on these cells the precise nature of this modulation is not clear. Further investigations into the long-term viability and/or recovery of cells following DTT treatment with and without PDGF were beyond the scope of this thesis. In future, such studies would provide further insight as to whether the addition of growth-factor is protective or harmful with respect to the response to DTT-induced stress.

7.2 DTT and UV-A stressors induce different changes in the ECM

7.2.1 DTT treated cells have decreased relative amounts of collagen proteins

A study by Reunanen et al., in 2000 used primary fibroblast cells from a healthy adult male donor and demonstrated that C_2-cermaide treatment reduced mRNA levels of collagen which could be reversed by the inhibition ERK1/2 and p38 MAPK [289]. Some years later, in 2010, another study by Kim et al., demonstrated the involvement of ERK1/2 activation on the IL-18 triggered downregulation of collagen at both the mRNA and protein level in skin fibroblasts extracted from new-born male foreskin[290]. In addition, STAT3 has recently been shown to enhance upregulation of collagen protein COL1A2. Activation of STAT3 is thought to allow the binding of STAT3 to the COL1A2 enhancer region and an increase in COL1A2 gene transcription[291]. Phosphorylation of STAT3 at Tyr705 is required for effective translocation to the nucleus where it can promote transcription of its target genes. There is some evidence to suggest that phosphorylation at Ser727 may impair subsequent phosphorylation of Tyr705 and this serine phosphorylation may be mediated by ERK-dependant or independent mechanisms[292].

In the investigations described in this thesis, reduced levels of collagen were seen with DTT treatment in BJ fibroblasts, a treatment that was also associated with dephosphorylation of p42/44 (ERK1/2) (Figure 4.12 and Figure 5.14). These investigations were carried out under serum free conditions which inactivate ERK signalling pathways[293]. This may explain the discrepancy in the connection between ERK phosphorylation and collagen production in the results presented in this thesis and those in the literature. It could, however imply that the reduction in collagen proteins seen in this data occurs independently of p42/44 activation.

Collagen proteins are characterised by the repeating “XaaYaaGly” motif, which runs along the central portion of the pro-collagen molecule. Most often amino acids X and Y are hydroxyproline and proline respectively and helical chains are formed from this repeating sequence. Three chains then assemble into a triple-helical procollagen fibril, bound at the N and C termini by two non-helical domains which are cleaved within the extracellular space to allow spontaneous assembly into collagen fibres[294]. Procollagen peptides are aligned before fibril formation in the ER by formation of intra- and inter-chain disulfide bonds at the C terminal pro-peptide domain[295]. Formation of the triple helix then proceeds from the C to N terminus. DTT treatment, which prohibits the formation of disulfide bonds within the ER, may therefore prevent correct formation of procollagen fibrils. Proteins that are unable to form the required disulfide bonds are targeted for degradation. This may provide an explanation for the reduction in collagen proteins seen following DTT treatment, which may be occurring independently of the signalling events detected.

This would also provide insight as to why a distinct reduction in relative collagen amount was not seen with UV-A radiation (Figure 6.9). Whilst UV-A radiation induces a redox stress response

associated with ROS generation following treatment, these ROS do not directly inhibit disulfide bond formation within the ER. Under these conditions, then, it is possible that correct formation of procollagen fibrils is still possible, resulting in the limited changes seen in relative collagen amounts. UV-damaged, or photo-aged skin is often associated with a decrease in the collagen ECM however evidence of this was not seen in the experimental model used for this thesis. It is likely that changes in collagen are, at least initially, mediated by an increased expression of collagenase enzymes including MMPs. Again, no evidence of increased relative amounts of MMPs was detected using this model system. It would be interesting, therefore, to assay the MMP activity in response to UV-A radiation to survey this further. It may also be important to identify the time-frame and UV-A radiation conditions under which collagen degradation occurs in these cells to fully investigate the specifics of their response.

7.2.2 UV-A radiation of cells promotes hyaluronan stability

Although no change in collagen proteins was detected in the UV-A treated sample data, there was some evidence of a change in the ECM environment that would promote stability of hyaluronan (see section 6.4). Hyaluronan, which is a glycosaminoglycan, is not possible to identify using a purely proteomic approach so the proteins that are involved in its synthesis and degradation were examined as a proxy. Although the hyaluronan synthesising enzymes, hyaluronan synthase 1,2 or 3 were not identified in the UV-A irradiated fibroblast data sets described here, there was a significant decrease in the relative levels of CEMIP, which mediates hyaluronan depolymerisation. In addition, there was a significant increase in inter-alpha-trypsin inhibitor (ITI) heavy chains H1, H2 and H3 which are known to bind hyaluronan, preventing its depolymerisation[272,296]. These changes are reflected in the significant overrepresentation of GO (biological process) terms 'negative regulation of endopeptidase activity', and 'hyaluronan metabolic process'.

Hyaluronan bound to ITI proteins has been detected in inflamed tissues[297,298]. In addition, hyaluronan has been shown to have antioxidant effects and promote Nrf2 activity within chondrocyte cells[299]. Chondrocytes are responsible for the secretion of ECM in cartilage, and the activity of Nrf2 transcription factor promotes transcription of antioxidant molecules and detoxifying enzymes to reduce the effects of oxidative stress[299]. The effects of UV-A radiation on hyaluronan specifically have not been studied. Rather, the published ECM-associated effects have focussed on regulation of collagenase enzymes and ECM protein degradation. However, it is possible that the hyaluronan stabilisation effects presented here may, in line with previous studies, be part of a UV-A induced inflammatory response, and cellular detoxification of ROS generated as a result of irradiation[300]. Repeating these experiments with chondrocytes may provide further insight into the regulation of hyaluronan stability and Nrf2 activity within a previously studied context.

Zoltan-Jones et al., demonstrated the ability of hyaluronan to drive the epithelial-mesenchymal transition (EMT) in normal epithelial cells[301] providing a link between hyaluronan and carcinogenesis, a well-known side-effect of excessive UV exposure. Prompted by this study, Jenkins et al., investigated the involvement of hyaluronan in the differentiation of fibroblasts to myofibroblasts[302]. Myofibroblasts are an activated class of fibroblast cell, typified by the apparently similar morphology to smooth muscle cells, and implicated in many diseased tissues including hypertrophic scar tissue and liver fibrosis[303]. Additionally, myofibroblasts have also been found associated with epithelial tumours, termed cancer-associated fibroblasts[303]. Jenkins et al., demonstrated that rather than driving differentiation (as was shown by Zoltan-Jones et al., with respect to the EMT), hyaluronan accumulated as a result of differentiation to myofibroblasts through a reduction in its turnover i.e. through stabilisation[302]. Although fibrosis, which is typically associated with excessive collagen deposition, is not associated with UV-A radiation, these studies pose the possibility that hyaluronan stabilisation may have more wide-reaching implications for differentiation of fibroblasts or stimulation of carcinogenesis. Given this potentially important role, and the indications of changes promoting hyaluronan stability found in the current data, it would be informative to investigate further the effect of UV-A radiation on hyaluronan by assaying for this non-sulfated glycosaminoglycan directly.

7.2.3 Comparison of the ECM response between DTT and UV-A radiation

The difference in responses seen between DTT treated and UV-A irradiated cells provides evidence for the specificity of the BJ fibroblast response to reductive stress, rather than a generalised response to redox imbalance. Treatment of cells with hydrogen peroxide induced signalling responses that were distinct from those seen with DTT treatment (Figure 4.11), and the responses of ECM proteins described above are distinct between DTT and UV-A treatments. Defined responses of the cells to oxidative versus reductive stress have been previously described for the ER chaperone protein BiP[304]. It therefore holds that there should be independent responses to reductive and oxidative stress that are distinct from a generalised redox stress response. Whilst much research has been done into the effects of oxidative stress, there is potential benefit to be harnessed from investigations into specific responses to reductive stress such as those described in this thesis. Although passive exposure to reductants is arguably less common than agents that induce an oxidative stress, it may be possible to actively use induction of mild reductive stresses to help in the treatment of certain disease states, such a fibrosis and the formation of excessive scar tissue.

7.3 SWATH as an investigative proteomic tool

This thesis has described the use of quantitative SWATH acquisition proteomics to profile the proteome of human dermal fibroblasts (BJ fibroblasts), and its response to both reducing (DTT) and oxidative (UV-A radiation) stress (Chapters 5 and 6 respectively). The ability of this technology to detect changes in secreted proteins from samples of conditioned media has also been tested, with some success.

7.3.1 Replicate variability

Perhaps not unsurprisingly, variability in SWATH quantification is higher in lower abundance peptides, namely those with a lower peak area[305]. In a study comparing the ability of SWATH to selected reaction monitoring (SRM), another popular technique in quantitative proteomics that is like SWATH but on a smaller scale, a low peak area was described as less than 1,000,000. In the DTT and PDGF treated sample data set approximately 10% of protein identifications were made with a summed peptide peak area of greater than 1,000,000. However, in the UV treated sample data set only 5% of protein identifications were made under this criterion, which may provide some explanation as to the increased variation seen in the UV data set. Here substantially fewer identifications met the %CV cut-off of 30% across all replicates compared to the number within this threshold from the DTT and PDGF data set.

The method used for SWATH acquisition described in this thesis used 100 variable Q1 isolation windows with 25 msec accumulation time, a standard protocol whereby narrower isolation windows are used for regions where peptide signals are most dense. A custom-designed protocol to determine the optimum isolation windows for each data set can be used and may improve identification and sample reproducibility in the results[306]. However, the 100 variable window method is largely accepted in the proteomic community (online[307] and through personal communication with Hubbard group, Manchester University, UK), and a trade-off must be made between minor gains to be had from custom-made window widths and the time and man-hours this takes on top of SWATH acquisition alone.

7.3.2 Secretome identifications

The use of SWATH acquisition afforded substantial benefit over DDA for the identification of proteins in conditioned media of cells, with over 700 proteins identified and quantified (see section 3.5). There was statistically significant overrepresentation of GO terms associated with the ECM confirming the detection of secreted proteins. Comparison of the secretome detected in DTT or etoposide treated cells had few statistically significant results, which may have been due to the low abundance of peptides detected. Only about 3% of protein identifications had summed peak area

of peptide peaks greater than 1,000,000 suggesting the peptide abundance was generally low. Some studies have used immunoaffinity peptide enrichment to increase abundance of low level peptides when looking at candidate biomarkers in patient serum or plasma[308,309]. These biomarkers are often only present at very low levels, but their presence is of clinical significance and the ability to employ MS improves sensitivity of assays compared to standard immunological techniques. It is possible that such a technique could be employed to mine for specific proteins of interest, for example to enrich conditioned media samples for MMP enzymes to detect changes that might occur in the relative abundance of these enzymes with DTT or UV-A treatment.

Concanavalin A purification was used in chapter 3 to enrich for secreted proteins in conditioned media before secretome analysis using a gel-based proteomic method. It is possible that returning to this technique and combining it with the increased power for SWATH MS would enable further detection of peptides, and confident identification of significantly changing proteins in the data.

7.3.3 Absolute vs relative quantification

Most large scale quantitative proteomic investigations employ relative quantification, comparing the signal intensities of peptides between samples from different treatment conditions. However, a drive from biomedical research, where it is beneficial to be able to quantify absolute amounts of biomarkers in patient samples, has led to the development of methods for absolute quantification. Schubert et al., developed a SWATH-based method for absolute quantification to investigate proteome dynamics in *Mycobacterium tuberculosis*[310]. This was based on a strategy first described by Ludwig et al., using SRM data, which details a linear correlation between the summed intensity of the top fragment ions from the most intense peptides and the absolute protein concentration[311]. Schubert et al., determined the correlation between signal intensities and absolute protein concentrations using spiked-in stable-isotope-labelled synthetic peptide standards and then applied this correlation to proteins identified in the SWATH data set.

Changes in relative amounts of enzyme may or may not have substantial impact on the activity of the cell depending on other variables such as the rate of turnover, and the basal expression level of that enzyme. The availability of absolute protein amounts allows the estimation of V_{max} values by multiplying absolute protein amounts with reported k_{cat} values of enzyme turnover. This data allows the investigation into the contribution of changes in enzyme amounts to the flux of metabolites within the cell[310].

It therefore may be of some value to measure the absolute protein concentrations during proteomic analysis, particularly when looking at enzymes. However, there is an additional cost from stable-isotope labelled standards for defining linear correlations used for absolute quantification and the more complex data analysis involved in interpreting this data. For the studies described in

this thesis, relative quantification by SWATH provides a good insight into the types of global proteomic changes that may be occurring in response to stress in these fibroblasts. These satisfy the aims of this thesis, which are not compromised by the lack of absolute quantification. Validation of the current findings could therefore be made by the use of absolute quantification, which wouldn't necessarily have to be done using label-free mass-spectrometry methodology. Alternatively, enzyme activity assays could accompany data to determine the relative impact of any change in relative protein amount, although these types of validations, while important, were beyond the scope of this thesis.

7.4 Considerations for future work

There are two main areas that could be considered in future proteomic investigations, following those described here.

First, the role of post-translational modifications in response to stress. It is well known that proteins may be modulated both in terms of their absolute abundance, as well as in the activation state of enzymes or targeting of proteins to specific subcellular compartments where they may or may not be able to function. Several studies have developed methods for the proteomic analysis of post-translational modifications, including those that occur in response to perturbations of the redox state for example glutathionylation[312–314]. The consideration of these modifications will add an extra layer to the proteomic investigations, allowing the functional analysis of a range of proteins that may not be regulated by changes in their absolute abundance. This may be of relevance to studies such as the one in this thesis, investigating the effect of redox modulating stress that may affect protein structures, for example with respect to the formation of disulfide bonds as was described for Akt at the beginning of this chapter.

Secondly, the culture conditions of the fibroblast cells. Fibroblasts, being the cells responsible for the secretion of the ECM, are normally resident amongst a complex network of ECM proteins, and not in densely packed monolayers as occur in the culture on 2D plastic culture flasks. Growth in a 3D environment, one example being the use of hydrogel scaffolds, may permit a more complete and representative formation of the ECM that may influence the cell behaviours observed. Additionally, it may also be possible to generate a model of the skin with multiple cell types to gain further insight in to the potential physiological relevance of the response to reductive or UV-A stress on a wider scale.

The investigations presented in this thesis provide good insight into the responses of fibroblasts to reductive stress, as well as providing a frame work through which to develop proteomic techniques with new quantitative capabilities. The inclusion of these factors would provide a fuller picture of the global proteomic response of dermal fibroblasts and the impact this may have on skin physiology as a whole.

7.5 Conclusions

This thesis has described several novel findings with respect to the responses of human dermal (BJ) fibroblasts to reductive stress and UV stress.

1. DTT stress induced a growth factor independent signalling response involving the chronic phosphorylation of Akt, and concomitant dephosphorylation of extracellular-signal regulated kinases p42 and p44.
2. PDGF modulated the cells global response to DTT treatment.
3. SWATH acquisition proteomics revealed a significant decrease in the relative amounts of several collagen proteins following DTT treatment.
4. Comparison of this response to oxidative stress induced by UV-A radiation suggests a response specific to reductive stress. This has been subject to limited specific investigations previously and may provide novel routes for therapeutic investigation.

Bibliography

Bibliography

1. Proksch, E., Brandner, J. M. & Jensen, J.-M. The skin: an indispensable barrier. *Experimental Dermatology* **17,** 1063–1072 (2008).

2. Eichenfield, L. F., Frieden, I. J., Zaenglein, A. & Mathes, E. *Neonatal and Infant Dermatology E-Book*. (Elsevier Health Sciences, 2014).

3. Moll, R., Divo, M. & Langbein, L. The human keratins: biology and pathology. *Histochem Cell Biol* **129,** 705 (2008).

4. Alberts, B. *et al. Molecular Biology of the Cell*. (Garland Science, 1994).

5. Watt, F. M., Jordan, P. W. & O'Neill, C. H. Cell shape controls terminal differentiation of human epidermal keratinocytes. *Proceedings of the National Academy of Sciences* **85,** 5576–5580 (1988).

6. Bikle, D. D., Xie, Z. & Tu, C.-L. Calcium regulation of keratinocyte differentiation. *Expert Rev Endocrinol Metab* **7,** 461–472 (2012).

7. Lowell, S., Jones, P., Le Roux, I., Dunne, J. & Watt, F. M. Stimulation of human epidermal differentiation by Delta–Notch signalling at the boundaries of stem-cell clusters. *Current Biology* **10,** 491–500 (2000).

8. Jones, P. H. & Watt, F. M. Separation of human epidermal stem cells from transit amplifying cells on the basis of differences in integrin function and expression. *Cell* **73,** 713–724 (1993).

9. Jones, P. H., Harper, S. & Watt, F. M. Stem cell patterning and fate in human epidermis. *Cell* **80,** 83–93 (1995).

10. Lang, D., Mascarenhas, J. B. & Shea, C. R. Melanocytes, melanocyte stem cells, and melanoma stem cells. *Clin Dermatol* **31,** 166–178 (2013).

11. Fajuyigbe, D. *et al.* Melanin distribution in human epidermis affords localized protection against DNA photodamage and concurs with skin cancer incidence difference in extreme phototypes. *The FASEB Journal* (2018).

12. Yamaguchi, Y. & Hearing, V. J. Melanocytes and Their Diseases. *Cold Spring Harb Perspect Med* **4,** (2014).

13. Yamaguchi, Y. & Hearing, V. J. Physiological factors that regulate skin pigmentation. *Biofactors* **35,** 193–199 (2009).

14. Bowman, S. L. & Marks, M. S. Shining a Light on Black Holes in Keratinocytes. *Journal of Investigative Dermatology* **138,** 486–489 (2018).

15. Losquadro, W. D. Anatomy of the Skin and the Pathogenesis of Nonmelanoma Skin Cancer. *Facial Plastic Surgery Clinics of North America* **25,** 283–289 (2017).

16. Romani, N., Brunner, P. M. & Stingl, G. Changing Views of the Role of Langerhans Cells. *Journal of Investigative Dermatology* **132,** 872–881 (2012).

17. Valladeau, J. *et al.* The monoclonal antibody DCGM4 recognizes Langerin, a protein specific of Langerhans cells, and is rapidly internalized from the cell surface. *Eur. J. Immunol.* **29,** 2695–2704 (1999).

18. Allan, R. S. *et al.* Epidermal viral immunity induced by CD8alpha+ dendritic cells but not by Langerhans cells. *Science* **301,** 1925–1928 (2003).

19. Kaplan, D. H., Jenison, M. C., Saeland, S., Shlomchik, W. D. & Shlomchik, M. J. Epidermal langerhans cell-deficient mice develop enhanced contact hypersensitivity. *Immunity* **23,** 611–620 (2005).

20. Kautz-Neu, K. *et al.* Langerhans cells are negative regulators of the anti-Leishmania response. *J Exp Med* **208,** 885–891 (2011).

21. Sandilands, A., Sutherland, C., Irvine, A. D. & McLean, W. H. I. Filaggrin in the frontline: role in skin barrier function and disease. *J Cell Sci* **122,** 1285–1294 (2009).

22. McLean, W. H. I. Filaggrin failure – from ichthyosis vulgaris to atopic eczema and beyond. *Br J Dermatol* **175,** 4–7 (2016).

23. Elias, P. M. & Feingold, K. R. *Skin Barrier*. (CRC Press, 2005).

24. Menon, G. K., Lee, S. E. & Lee, S. H. An Overview of Epidermal Lamellar Bodies: Novel roles in biological adaptations and secondary barriers. *Journal of Dermatological Science* (2018). doi:10.1016/j.jdermsci.2018.03.005

25. Rice, R. H. & Green, H. The cornified envelope of terminally differentiated human epidermal keratinocytes consists of cross-linked protein. *Cell* **11,** 417–422 (1977).

26. Kalinin, A. E., Kajava, A. V. & Steinert, P. M. Epithelial barrier function: assembly and structural features of the cornified cell envelope. *BioEssays* **24,** 789–800 (2002).

27. Garrod, D. & Chidgey, M. Desmosome structure, composition and function. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1778,** 572–587 (2008).

28. Hovnanian, A. Netherton syndrome: skin inflammation and allergy by loss of protease inhibition. *Cell Tissue Res.* **351,** 289–300 (2013).

29. Johnson, J. L., Najor, N. A. & Green, K. J. Desmosomes: Regulators of Cellular Signaling and Adhesion in Epidermal Health and Disease. *Cold Spring Harb Perspect Med* **4,** (2014).

30. Li, S. *et al.* Laminin–sulfatide binding initiates basement membrane assembly and enables receptor signaling in Schwann cells and fibroblasts. *J Cell Biol* **169,** 179–189 (2005).

31. Yurchenco, P. D. Basement Membranes: Cell Scaffoldings and Signaling Platforms. *Cold Spring Harb Perspect Biol* **3,** (2011).

32. Sasaki, T., Fässler, R. & Hohenester, E. Laminin. *J Cell Biol* **164,** 959–963 (2004).

33. Durbeej, M. Laminins. *Cell Tissue Res* **339,** 259 (2010).

34. Ho, M. S. P., Böse, K., Mokkapati, S., Nischt, R. & Smyth, N. Nidogens-Extracellular matrix linker molecules. *Microsc. Res. Tech.* **71,** 387–395 (2008).

35. Paavola, K. J., Sidik, H., Zuchero, J. B., Eckart, M. & Talbot, W. S. Type IV collagen is an activating ligand for the adhesion G protein-coupled receptor GPR126. *Sci Signal* **7,** ra76 (2014).

36. Stawikowski, M. J., Aukszi, B., Stawikowska, R., Cudic, M. & Fields, G. B. Glycosylation Modulates Melanoma Cell α2β1 and α3β1 Integrin Interactions with Type IV Collagen. *J Biol Chem* **289,** 21591–21604 (2014).

37. Petitclerc, E. *et al.* New Functions for Non-collagenous Domains of Human Collagen Type IV NOVEL INTEGRIN LIGANDS INHIBITING ANGIOGENESIS AND TUMOR GROWTHIN VIVO. *J. Biol. Chem.* **275,** 8051–8061 (2000).

38. Ikeda, K. *et al.* Loss of Expression of Type IV Collagen α5 and α6 Chains in Colorectal Cancer Associated with the Hypermethylation of Their Promoter Region. *Am J Pathol* **168,** 856–865 (2006).

39. Walko, G., Castañón, M. J. & Wiche, G. Molecular architecture and function of the hemidesmosome. *Cell Tissue Res* **360,** 363–378 (2015).

40. Sorrell, J. M. & Caplan, A. I. Fibroblast heterogeneity: more than skin deep. *J. Cell. Sci.* **117,** 667–675 (2004).

41. Thangapazham, R. L., Darling, T. N. & Meyerle, J. Alteration of Skin Properties with Autologous Dermal Fibroblasts. *Int J Mol Sci* **15,** 8407–8427 (2014).

42. Janson, D. G., Saintigny, G., van Adrichem, A., Mahé, C. & El Ghalbzouri, A. Different gene expression patterns in human papillary and reticular fibroblasts. *J. Invest. Dermatol.* **132,** 2565–2572 (2012).

43. Sorrell, J. M., Baber, M. A. & Caplan, A. I. Human dermal fibroblast subpopulations; differential interactions with vascular endothelial cells in coculture: nonsoluble factors in the extracellular matrix influence interactions. *Wound Repair Regen* **16,** 300–309 (2008).

44. Harper, R. A. & Grove, G. Human skin fibroblasts derived from papillary and reticular dermis: differences in growth potential in vitro. *Science* **204,** 526–527 (1979).

45. Schafer, I. A., Pandy, M., Ferguson, R. & Davis, B. R. Comparative observation of fibroblasts derived from the papillary and reticular dermis of infants and adults: growth kinetics, packing density at confluence and surface morphology. *Mech. Ageing Dev.* **31,** 275–293 (1985).

46. Driskell, R. R. *et al.* Distinct fibroblast lineages determine dermal architecture in skin development and repair. *Nature* **504,** 277–281 (2013).

47. Driskell, R. R. & Watt, F. M. Understanding fibroblast heterogeneity in the skin. *Trends Cell Biol.* **25,** 92–99 (2015).

48. Kretzschmar, K. *et al.* BLIMP1 is required for postnatal epidermal homeostasis but does not define a sebaceous gland progenitor under steady-state conditions. *Stem Cell Reports* **3,** 620–633 (2014).

49. Rinkevich, Y. *et al.* Skin fibrosis. Identification and isolation of a dermal lineage with intrinsic fibrogenic potential. *Science* **348,** aaa2151 (2015).

50. Driskell, R., Jahoda, C. A. B., Chuong, C.-M., Watt, F. & Horsley, V. Defining dermal adipose tissue. *Exp Dermatol* **23,** 629–631 (2014).

51. Kelley, D. E., Thaete, F. L., Troost, F., Huwe, T. & Goodpaster, B. H. Subdivisions of subcutaneous abdominal adipose tissue and insulin resistance. *American Journal of Physiology-Endocrinology and Metabolism* **278,** E941–E948 (2000).

52. Schmidt, B. A. & Horsley, V. Intradermal adipocytes mediate fibroblast recruitment during skin wound healing. *Development* **140,** 1517–1527 (2013).

53. Alexander, C. M. *et al.* Dermal white adipose tissue: a new component of the thermogenic response. *J Lipid Res* **56,** 2061–2069 (2015).

54. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181,** 223–230 (1973).

55. Levinthal, C. J. Are there pathways for protein folding? *Journal de Chimie Physique* **65,** 44–45 (1968).

56. Kamtekar, S., Schiffer, J. M., Xiong, H., Babik, J. M. & Hecht, M. H. Protein design by binary patterning of polar and nonpolar amino acids. *Science* **262,** 1680–1685 (1993).

57. Dill, K. A., Ozkan, S. B., Shell, M. S. & Weikl, T. R. The Protein Folding Problem. *Annu Rev Biophys* **37,** 289–316 (2008).

58. Kim, Y. E., Hipp, M. S., Bracher, A., Hayer-Hartl, M. & Hartl, F. U. Molecular Chaperone Functions in Protein Folding and Proteostasis. *Annual Review of Biochemistry* **82,** 323–355 (2013).

59. Dobson, C. M. Protein folding and misfolding. *Nature* (2003). doi:10.1038/nature02261

60. EGF-like domain (IPR000742) < InterPro < EMBL-EBI. Available at: https://www.ebi.ac.uk/interpro/entry/IPR000742. (Accessed: 18th May 2018)

61. Campbell, I. D. & Bork, P. Epidermal growth factor-like modules. *Current Opinion in Structural Biology* **3,** 385–392 (1993).

62. Ellgaard, L. Catalysis of disulphide bond formation in the endoplasmic reticulum. *Biochem. Soc. Trans.* **32,** 663–667 (2004).

63. Robinson, P. J., Pringle, M. A., Woolhead, C. A. & Bulleid, N. J. Folding of a single domain protein entering the endoplasmic reticulum precedes disulfide formation. *J Biol Chem* **292,** 6978–6986 (2017).

64. Benham, A. M. Protein Secretion and the Endoplasmic Reticulum. *Cold Spring Harb Perspect Biol* **4,** a012872 (2012).

65. von Heijne, G. Signal sequences. The limits of variation. *J. Mol. Biol.* **184,** 99–105 (1985).

66. Lakkaraju, A. K. K., Mary, C., Scherrer, A., Johnson, A. E. & Strub, K. SRP Keeps Polypeptides Translocation-Competent by Slowing Translation to Match Limiting ER-Targeting Sites. *Cell* **133,** 440–451 (2008).

67. Guerriero, C. J. & Brodsky, J. L. The delicate balance between secreted protein folding and Endoplasmic Reticulum-associated degradation in human physiology. *Physiol Rev* **92,** 537–576 (2012).

68. Netzer, W. J. & Hartl, F. U. Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* **388,** 343–349 (1997).

69. Buck, T. M., Wright, C. M. & Brodsky, J. L. The Activities and Function of Molecular Chaperones in the Endoplasmic Reticulum. *Semin Cell Dev Biol* **18,** 751–761 (2007).

70. Mayer, M. P. & Bukau, B. Hsp70 chaperones: Cellular functions and molecular mechanism. *CMLS, Cell. Mol. Life Sci.* **62,** 670 (2005).

71. Matlack, K. E., Misselwitz, B., Plath, K. & Rapoport, T. A. BiP acts as a molecular ratchet during posttranslational transport of prepro-alpha factor across the ER membrane. *Cell* **97,** 553–564 (1999).

72. Whitesell, L. & Lindquist, S. L. HSP90 and the chaperoning of cancer. *Nature Reviews Cancer* **5,** 761–772 (2005).

73. Eletto, D., Dersh, D. & Argon, Y. GRP94 in ER Quality Control and Stress Responses. *Semin Cell Dev Biol* **21,** 479–485 (2010).

74. Morán Luengo, T., Kityk, R., Mayer, M. P. & Rüdiger, S. G. D. Hsp90 Breaks the Deadlock of the Hsp70 Chaperone System. *Molecular Cell* **70,** 545-552.e9 (2018).

75. Trombetta, E. S. & Helenius, A. Lectins as chaperones in glycoprotein folding. *Current Opinion in Structural Biology* **8,** 587–592 (1998).

76. Trombetta, E. S. & Helenius, A. Conformational Requirements for Glycoprotein Reglucosylation in the Endoplasmic Reticulum. *The Journal of Cell Biology* **148,** 1123–1130 (2000).

77. Tasab, M., Batten, M. R. & Bulleid, N. J. Hsp47: a molecular chaperone that interacts with and stabilizes correctly-folded procollagen. *The EMBO Journal* **19,** 2204–2211 (2000).

78. Nagata, K. Hsp47: a collagen-specific molecular chaperone. *Trends in Biochemical Sciences* **21,** 23–26 (1996).

79. Ryhänen, L., Zaragoza, E. J. & Uitto, J. Conformational stability of type I collagen triple helix: Evidence for temporary and local relaxation of the protein conformation using a proteolytic probe. *Archives of Biochemistry and Biophysics* **223,** 562–571 (1983).

80. Duran, I. *et al.* HSP47 and FKBP65 cooperate in the synthesis of type I procollagen. *Hum Mol Genet* **24,** 1918–1928 (2015).

81. Bulleid, N. J. Disulfide Bond Formation in the Mammalian Endoplasmic Reticulum. *Cold Spring Harb Perspect Biol* **4,** (2012).

82. Mor-Cohen, R. Disulfide Bonds as Regulators of Integrin Function in Thrombosis and Hemostasis. *Antioxidants & Redox Signaling* **24,** 16–31 (2014).

83. Jansens, A., van Duijn, E. & Braakman, I. Coordinated nonvectorial folding in a newly synthesized multidomain protein. *Science* **298,** 2401–2403 (2002).

84. Tavender, T. J., Springate, J. J. & Bulleid, N. J. Recycling of peroxiredoxin IV provides a novel pathway for disulphide formation in the endoplasmic reticulum. *EMBO J* **29,** 4185–4197 (2010).

85. Braakman, I. & Hebert, D. N. Protein Folding in the Endoplasmic Reticulum. *Cold Spring Harb Perspect Biol* **5,** (2013).

86. Matena, A., Rehic, E., Hönig, D., Kamba, B. & Bayer, P. Structure and function of the human parvulins Pin1 and Par14/17. *Biological Chemistry* **399,** 101–125 (2018).

87. Bächinger, H. P. The influence of peptidyl-prolyl cis-trans isomerase on the in vitro folding of type III collagen. *J. Biol. Chem.* **262,** 17144–17148 (1987).

88. Steinmann, B., Bruckner, P. & Superti-Furga, A. Cyclosporin A slows collagen triple-helix formation in vivo: indirect evidence for a physiologic role of peptidyl-prolyl cis-trans-isomerase. *J. Biol. Chem.* **266,** 1299–1303 (1991).

89. Chen, Y. *et al.* FKBP65-dependent peptidyl-prolyl isomerase activity potentiates the lysyl hydroxylase 2-driven collagen cross-link switch. *Scientific Reports* **7,** 46021 (2017).

90. Ruggiano, A., Foresti, O. & Carvalho, P. ER-associated degradation: Protein quality control and beyond. *J Cell Biol* **204,** 869–879 (2014).

91. Schubert, U. *et al.* Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* **404,** 770–774 (2000).

92. Ellgaard, L. & Helenius, A. Quality control in the endoplasmic reticulum. *Nature Reviews Molecular Cell Biology* **4,** 181–191 (2003).

93. Frenkel, Z., Gregory, W., Kornfeld, S. & Lederkremer, G. Z. Endoplasmic reticulum-associated degradation of mammalian glycoproteins involves sugar chain trimming to Man6-5GlcNAc2. *J. Biol. Chem.* **278,** 34119–34124 (2003).

94. Słomińska-Wojewódzka, M. & Sandvig, K. The Role of Lectin-Carbohydrate Interactions in the Regulation of ER-Associated Protein Degradation. *Molecules* **20,** 9816–9846 (2015).

95. Cormier, J. H., Tamura, T., Sunryd, J. C. & Hebert, D. N. EDEM1 recognition and delivery of misfolded proteins to the SEL1L-containing ERAD complex. *Mol Cell* **34,** 627–633 (2009).

96. Herrmann, J. M., Malkus, P. & Schekman, R. Out of the ER—outfitters, escorts and guides. *Trends in Cell Biology* **9,** 5–7 (1999).

97. Song, B.-L., Sever, N. & DeBose-Boyd, R. A. Gp78, a membrane-anchored ubiquitin ligase, associates with Insig-1 and couples sterol-regulated ubiquitination to degradation of HMG CoA reductase. *Mol. Cell* **19,** 829–840 (2005).

98. Schubert, U. *et al.* CD4 Glycoprotein Degradation Induced by Human Immunodeficiency Virus Type 1 Vpu Protein Requires the Function of Proteasomes and the Ubiquitin-Conjugating Pathway. *J. Virol.* **72,** 2280–2288 (1998).

99. Sommer, T. & Jentsch, S. A protein translocation defect linked to ubiquitin conjugation at the endoplasmic reticulum. *Nature* **365,** 176–179 (1993).

100. Ward, C. L., Omura, S. & Kopito, R. R. Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell* **83,** 121–127 (1995).

101. Taxis, C. *et al.* Use of Modular Substrates Demonstrates Mechanistic Diversity and Reveals Differences in Chaperone Requirement of ERAD. *J. Biol. Chem.* **278,** 35903–35913 (2003).

102. Wu, X. & Rapoport, T. A. Mechanistic insights into ER-associated protein degradation. *Current Opinion in Cell Biology* **53,** 22–28 (2018).

103. Baldridge, R. D. & Rapoport, T. A. Autoubiquitination of the Hrd1 Ligase Triggers Protein Retrotranslocation in ERAD. *Cell* **166,** 394–407 (2016).

104. Fregno, I. & Molinari, M. Endoplasmic reticulum turnover: ER-phagy and other flavors in selective and non-selective ER clearance. *F1000Res* **7,** (2018).

105. Schröder, M. & Kaufman, R. J. ER stress and the unfolded protein response. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **569,** 29–63 (2005).

106. Walter, P. & Ron, D. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. *Science* **334,** 1081–1086 (2011).

107. Gardner, B. M., Pincus, D., Gotthardt, K., Gallagher, C. M. & Walter, P. Endoplasmic Reticulum Stress Sensing in the Unfolded Protein Response. *Cold Spring Harb Perspect Biol* **5,** a013169 (2013).

108. He, Y. *et al.* Emerging roles for XBP1, a sUPeR transcription factor. *Gene Expr* **15,** 13–25 (2010).

109. Tam, A. B., Koong, A. C. & Niwa, M. Ire1 Has Distinct Catalytic Mechanisms for XBP1/HAC1 Splicing and RIDD. *Cell Reports* **9,** 850–858 (2014).

110. Jackson, R. J., Hellen, C. U. T. & Pestova, T. V. THE MECHANISM OF EUKARYOTIC TRANSLATION INITIATION AND PRINCIPLES OF ITS REGULATION. *Nat Rev Mol Cell Biol* **11,** 113–127 (2010).

111. Veal, E. A., Toone, W. M., Jones, N. & Morgan, B. A. Distinct Roles for Glutathione S-Transferases in the Oxidative Stress Response in Schizosaccharomyces pombe. *J. Biol. Chem.* **277,** 35523–35531 (2002).

112. Cullinan, S. B. & Diehl, J. A. Coordination of ER and oxidative stress signaling: The PERK/Nrf2 signaling pathway. *The International Journal of Biochemistry & Cell Biology* **38,** 317–332 (2006).

113. Anti-aging Market- Global Industry Analysis, Trends, Size, Forecast 2013 - 2019. Available at: https://www.transparencymarketresearch.com/anti-aging-market.html. (Accessed: 24th May 2018)

114. Vonk, L. A. *et al.* Endoplasmic reticulum stress inhibits collagen synthesis independent of collagen-modifying enzymes in different chondrocyte populations and dermal fibroblasts. *Biochem. Cell Biol.* **88,** 539–552 (2010).

115. Chong, W. C., Shastri, M. D. & Eri, R. Endoplasmic Reticulum Stress and Oxidative Stress: A Vicious Nexus Implicated in Bowel Disease Pathophysiology. *Int J Mol Sci* **18,** (2017).

116. Cao, S. S. & Kaufman, R. J. Endoplasmic Reticulum Stress and Oxidative Stress in Cell Fate Decision and Human Disease. *Antioxid Redox Signal* **21,** 396–413 (2014).

117. van der Vlies, D., Wirtz, K. W. & Pap, E. H. Detection of protein oxidation in rat-1 fibroblasts by fluorescently labeled tyramine. *Biochemistry* **40,** 7783–7788 (2001).

118. Vlies, D. van der, Pap, E. H. W., Post, J. A., Celis, J. E. & Wirtz, K. W. A. Endoplasmic reticulum resident proteins of normal human dermal fibroblasts are the major targets for oxidative stress induced by hydrogen peroxide. *Biochemical Journal* **366,** 825–830 (2002).

119. Harman, D. The aging process. *PNAS* **78,** 7124–7128 (1981).

120. Fisher, G. J. *et al.* Collagen Fragmentation Promotes Oxidative Stress and Elevates Matrix Metalloproteinase-1 in Fibroblasts in Aged Human Skin. *The American Journal of Pathology* **174,** 101–114 (2009).

121. Trouba, K. J., Hamadeh, H. K., Amin, R. P. & Germolec, D. R. Oxidative Stress and Its Role in Skin Disease. *Antioxidants & Redox Signaling* **4,** 665–673 (2002).

122. Komori, R. *et al.* Ultraviolet a induces endoplasmic reticulum stress response in human dermal fibroblasts. *Cell Struct. Funct.* **37,** 49–53 (2012).

123. Wondrak, G. T., Roberts, M. J., Cervantes-Laurean, D., Jacobson, M. K. & Jacobson, E. L. Proteins of the Extracellular Matrix Are Sensitizers of Photo-oxidative Stress in Human Skin Cells. *Journal of Investigative Dermatology* **121,** 578–586 (2003).

124. Calabrese, V. *et al.* Redox regulation of cellular stress response by ferulic acid ethyl ester in human dermal fibroblasts: role of vitagenes. *Clinics in Dermatology* **26,** 358–363 (2008).

125. Giampieri, F. *et al.* Polyphenol-Rich Strawberry Extract Protects Human Dermal Fibroblasts against Hydrogen Peroxide Oxidative Damage and Improves Mitochondrial Functionality. *Molecules* **19,** 7798–7816 (2014).

126. Wu, C.-L. *et al.* Proteomic analysis of UVB-induced protein expression- and redox-dependent changes in skin fibroblasts using lysine- and cysteine-labeling two-dimensional difference gel electrophoresis. *Journal of Proteomics* **75,** 1991–2014 (2012).

127. Griffin, J. L. The Cinderella story of metabolic profiling: does metabolomics get to go to the functional genomics ball? *Philos Trans R Soc Lond B Biol Sci* **361,** 147–161 (2006).

128. Tyers, M. & Mann, M. From genomics to proteomics. *Nature* (2003). doi:10.1038/nature01510

129. Liebler, D. C. *Introduction to Proteomics: Tools for the New Biology*. (Springer Science & Business Media, 2001).

130. Cañas, B., López-Ferrer, D., Ramos-Fernández, A., Camafeita, E. & Calvo, E. Mass spectrometry technologies for proteomics. *Brief Funct Genomics* **4,** 295–320 (2006).

131. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246,** 64–71 (1989).

132. Karas, M. & Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **60,** 2299–2301 (1988).

133. Bruins, A. P. Mechanistic aspects of electrospray ionization. *Journal of Chromatography A* **794,** 345–357 (1998).

134. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422,** 198–207 (2003).

135. Lovric, J. *Introducing proteomics from concepts to sample separation, mass spectrometry and data analysis*. (Wiley-Blackwell, 2011).

136. Nilsson, T. *et al.* Mass spectrometry in high-throughput proteomics: ready for the big time. *Nature Methods* (2010). doi:10.1038/nmeth0910-681

137. Edelmann, M. J. Strong Cation Exchange Chromatography in Analysis of Posttranslational Modifications: Innovations and Perspectives. *BioMed Research International* (2011). doi:10.1155/2011/936508

138. Benham, A. M. *et al.* The CXXCXXC motif determines the folding, structure and stability of human Ero1-Lα. *EMBO J* **19,** 4493–4502 (2000).

139. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2,** 1896–1906 (2007).

140. Smith, S. J., Kroon, J. T. M., Simon, W. J., Slabas, A. R. & Chivasa, S. A Novel Function for *Arabidopsis* CYCLASE1 in Programmed Cell Death Revealed by Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) Analysis of Extracellular Matrix Proteins. *Molecular & Cellular Proteomics* **14,** 1556–1568 (2015).

141. Rosenberger, G. *et al.* A repository of assays to quantify 10,000 human proteins by SWATH-MS. *Sci Data* **1,** 140031 (2014).

142. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology* **16,** 284–287 (2012).

143. Mi, H. *et al.* PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res* **45,** D183–D189 (2017).

144. Thomas, P. D. *et al.* Applications for protein sequence–function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res* **34,** W645–W650 (2006).

145. Chernushevich, I. V., Loboda, A. V. & Thomson, B. A. An introduction to quadrupole–time-of-flight mass spectrometry. *J. Mass Spectrom.* **36,** 849–865 (2001).

146. Gillet, L. C. *et al.* Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Mol Cell Proteomics* **11,** (2012).

147. Keller, B. O., Sui, J., Young, A. B. & Whittal, R. M. Interferences and contaminants encountered in modern mass spectrometry. *Analytica Chimica Acta* **627,** 71–81 (2008).

148. Mirzaei, H. & Carrasco, M. *Modern Proteomics – Sample Preparation, Analysis and Practical Applications*. (Springer, 2016).

149. Chae, P. S. *et al.* Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat. Methods* **7,** 1003–1008 (2010).

150. Seddon, A. M., Curnow, P. & Booth, P. J. Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1666,** 105–117 (2004).

151. Brown, R. B. & Audet, J. Current techniques for single-cell lysis. *J R Soc Interface* **5,** S131–S138 (2008).

152. Kachuk, C., Faulkner, M., Liu, F. & Doucette, A. A. Automated SDS Depletion for Mass Spectrometry of Intact Membrane Proteins though Transmembrane Electrophoresis. *J. Proteome Res.* **15,** 2634–2642 (2016).

153. Benham, A. M., van Lith, M., Sitia, R. & Braakman, I. Ero1-PDI interactions, the response to redox flux and the implications for disulfide bond formation in the mammalian endoplasmic reticulum. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **368,** 20110403 (2013).

154. Yang, L., Zhang, H. & Bruce, J. E. Optimizing the detergent concentration conditions for immunoprecipitation (IP) coupled with LC-MS/MS identification of interacting proteins. *The Analyst* **134,** 755 (2009).

155. Scheraga, H. A. The thrombin–fibrinogen interaction. *Biophysical Chemistry* **112,** 117–130 (2004).

156. Hahm, E. *et al.* Extracellular protein disulfide isomerase regulates ligand-binding activity of αMβ2 integrin and neutrophil recruitment during vascular inflammation. *Blood* **121,** 3789–3800 (2013).

157. Reinhardt, C. *et al.* Protein disulfide isomerase acts as an injury response signal that enhances fibrin generation via tissue factor activation. *J Clin Invest* **118,** 1110–1122 (2008).

158. Stopa, J. D. *et al.* Protein disulfide isomerase inhibition blocks thrombin generation in humans by interfering with platelet factor V activation. *JCI Insight* **2,** (2017).

159. Swiatkowska, M. *et al.* Ero1α Is Expressed on Blood Platelets in Association with Protein-disulfide Isomerase and Contributes to Redox-controlled Remodeling of αIIbβ3. *J. Biol. Chem.* **285,** 29874–29883 (2010).

160. Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43,** D447-452 (2015).

161. Quan, C. *et al.* Dermal fibroblast expression of stromal cell-derived factor-1 (SDF-1) promotes epidermal keratinocyte proliferation in normal and diseased skin. *Protein & Cell* **6,** 890 (2015).

162. BJ ATCC ® CRL-2522™ Homo sapiens skin; foreskin normal. Available at: https://www.lgcstandards-atcc.org/Products/Cells_and_Microorganisms/By_Focus_Area/Toxicology/CRL-2522.aspx?geo_country=gb#characteristics. (Accessed: 5th June 2018)

163. Coppé, J.-P., Desprez, P.-Y., Krtolica, A. & Campisi, J. The Senescence-Associated Secretory Phenotype: The Dark Side of Tumor Suppression. *Annu Rev Pathol* **5,** 99–118 (2010).

164. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nature Methods* **6,** 359–362 (2009).

165. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology* **16,** 284–287 (2012).

166. Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nature Protocols* **8,** 1551–1566 (2013).

167. Benjamini, Y. & Hochberg, Y. A direct approach to false discovery rates. *JR Stat. Soc.* **57,** 289–300 (1995).

168. McDonald, J. H. *Handbook of Biological Statistics*. (Sparky House Publishing, Baltimore, Maryland).

169. Deppermann, C. *et al.* Gray platelet syndrome and defective thrombo-inflammation in *Nbeal2*-deficient mice. *J Clin Invest* **123,** 3331–3342 (2013).

170. Neilson Karlie A. *et al.* Less label, more free: Approaches in label-free quantitative mass spectrometry. *PROTEOMICS* **11,** 535–553 (2011).

171. Patel, V. J. *et al.* A Comparison of Labeling and Label-Free Mass Spectrometry-Based Proteomics Approaches. *J. Proteome Res.* **8,** 3752–3759 (2009).

172. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. & Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* **389,** 1017–1031 (2007).

173. Gos, M. *et al.* Cellular quiescence induced by contact inhibition or serum withdrawal in C3H10T1/2 cells. *Cell Prolif.* **38,** 107–116 (2005).

174. Hayes, O. *et al.* Cell confluency is as efficient as serum starvation for inducing arrest in the G0/G1 phase of the cell cycle in granulosa and fibroblast cells of cattle. *Anim. Reprod. Sci.* **87,** 181–192 (2005).

175. Coller, H. A., Sang, L. & Roberts, J. M. A New Description of Cellular Quiescence. *PLOS Biology* **4,** e83 (2006).

176. Nishiyama, T. *et al.* Response to growth factors of human dermal fibroblasts in a quiescent state owing to cell-matrix contact inhibition. *Matrix* **11,** 71–75 (1991).

177. Ross, R., Glomset, J., Kariya, B. & Harker, L. A Platelet-Dependent Serum Factor That Stimulates the Proliferation of Arterial Smooth Muscle Cells In Vitro. *Proceedings of the National Academy of Sciences of the United States of America* **71,** 1207–1210 (1974).

178. Chen, P.-H., Chen, X. & He, X. Platelet-derived growth factors and their receptors: structural and functional perspectives. *Biochim Biophys Acta* **1834,** 2176–2186 (2013).

179. Heldin, C.-H. & Westermark, B. Mechanism of Action and In Vivo Role of Platelet-Derived Growth Factor. *Physiological Reviews* **79,** 1283–1316 (1999).

180. Reigstad, L. J., Varhaug, J. E. & Lillehaug, J. R. Structural and functional specificities of PDGF-C and PDGF-D, the novel members of the platelet-derived growth factors family. *The FEBS Journal* **272,** 5723–5741 (2005).

181. Demoulin, J.-B. & Essaghir, A. PDGF receptor signaling networks in normal and cancer cells. *Cytokine & Growth Factor Reviews* **25,** 273–283 (2014).

182. Andrae, J., Gallini, R. & Betsholtz, C. Role of platelet-derived growth factors in physiology and medicine. *Genes & Development* **22,** 1276–1312 (2008).

183. Iwayama, T. & Olson, L. E. Involvement of PDGF in Fibrosis and Scleroderma: Recent Insights from Animal Models and Potential Therapeutic Opportunities. *Curr Rheumatol Rep* **15,** 304 (2013).

184. Distler, J. H. W. *et al.* Imatinib mesylate reduces production of extracellular matrix and prevents development of experimental dermal fibrosis. *Arthritis Rheum.* **56,** 311–322 (2007).

185. Gabrielli, A., Svegliati, S., Moroncini, G. & Amico, D. New Insights into the Role of Oxidative Stress in Scleroderma Fibrosis. *Open Rheumatol J* **6,** 87–95 (2012).

186. Cleland, W. W. Dithiothreitol, a New Protective Reagent for SH Groups*. *Biochemistry* **3,** 480–482 (1964).

187. Rand, J. D. & Grant, C. M. The Thioredoxin System Protects Ribosomes against Stress-induced Aggregation. *Mol Biol Cell* **17,** 387–401 (2006).

188. Bertolotti, A., Zhang, Y., Hendershot, L. M., Harding, H. P. & Ron, D. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nature Cell Biology* **2,** 326–332 (2000).

189. Calfon, M. *et al.* IRE1 couples endoplasmic reticulum load to secretory capacity by processing the *XBP-1* mRNA. *Nature* **415,** 92–96 (2002).

190. Lemin Andrew J., Saleki Khalil, van Lith Marcel & Benham Adam M. Activation of the unfolded protein response and alternative splicing of ATF6α in HLA-B27 positive lymphocytes. *FEBS Letters* **581,** 1819–1824 (2007).

191. Hinnebusch, A. G. The eIF-2α kinases: regulators of protein synthesis in starvation and stress. *Seminars in Cell Biology* **5,** 417–426 (1994).

192. Manning, B. D. & Cantley, L. C. AKT/PKB Signaling: Navigating Downstream. *Cell* **129,** 1261–1274 (2007).

193. Mebratu, Y. & Tesfaigzi, Y. How ERK1/2 Activation Controls Cell Proliferation and Cell Death Is Subcellular Localization the Answer? *Cell Cycle* **8,** 1168–1175 (2009).

194. Krebs, J., Agellon, L. B. & Michalak, M. Ca2+ homeostasis and endoplasmic reticulum (ER) stress: An integrated view of calcium signaling. *Biochemical and Biophysical Research Communications* **460,** 114–121 (2015).

195. Collin, C. *et al.* Protective effects of taurine on human hair follicle grown in vitro1. *International Journal of Cosmetic Science* **28,** 289–298 (2006).

196. Blaisdell, R. J. & Giri, S. N. Mechanism of antifibrotic effect of taurine and niacin in the multidose bleomycin-hamster model of lung fibrosis: inhibition of lysyl oxidase and collagenase. *J. Biochem. Toxicol.* **10,** 203–210 (1995).

197. Wang, Q., Giri, S. N., Hyde, D. M. & Nakashima, J. M. Effects of Taurine on Bleomycin-Induced Lung Fibrosis in Hamsters. *Proceedings of the Society for Experimental Biology and Medicine* **190,** 330–338 (1989).

198. Trachtman, H., Futterweit, S. & Bienkowski, R. S. Taurine Prevents Glucose-Induced Lipid Peroxidation and Increased Collagen Production in Cultured Rat Mesangial Cells. *Biochemical and Biophysical Research Communications* **191,** 759–765 (1993).

199. Pancar, G. S. & Kalkan, G. Irritant nail dermatitis of chemical depilatory product presenting with koilonychia. *Cutan Ocul Toxicol* **33,** 87–89 (2014).

200. Park, J. H. *et al.* Akt attenuates apoptotic death through phosphorylation of H2A under hydrogen peroxide-induced oxidative stress in PC12 cells and hippocampal neurons. *Sci Rep* **6,** (2016).

201. Crossthwaite, A. J., Hasan, S. & Williams, R. J. Hydrogen peroxide-mediated phosphorylation of ERK1/2, Akt/PKB and JNK in cortical neurones: dependence on Ca2+ and PI3-kinase. *Journal of Neurochemistry* **80,** 24–35 (2002).

202. Jong, C. J., Azuma, J. & Schaffer, S. Mechanism underlying the antioxidant activity of taurine: prevention of mitochondrial oxidant production. *Amino Acids* **42,** 2223–2232 (2012).

203. Marcinkiewicz, J. & Schaffer, S. W. *Taurine 9*. (Springer, 2015).

204. Richardson, P. G., Eng, C., Kolesar, J., Hideshima, T. & Anderson, K. C. Perifosine, an oral, anti-cancer agent and inhibitor of the Akt pathway: mechanistic actions, pharmacodynamics, pharmacokinetics and clinical activity. *Expert Opinion on Drug Metabolism & Toxicology* **8,** 623–633 (2012).

205. Hideshima, T. *et al.* Perifosine, an oral bioactive novel alkylphospholipid, inhibits Akt and induces in vitro and in vivo cytotoxicity in human multiple myeloma cells | Blood Journal. *Blood* **107,** 4053–4062 (2006).

206. Astle, M. V. *et al.* AKT induces senescence in human cells via mTORC1 and p53 in the absence of DNA damage: implications for targeting mTOR during malignancy. *Oncogene* **31,** 1949–1962 (2012).

207. Nogueira, V. *et al.* Akt determines replicative senescence and oxidative or oncogenic premature senescence and sensitizes cells to oxidative apoptosis. *Cancer Cell* **14,** 458–470 (2008).

208. Bent, E. H., Gilbert, L. A. & Hemann, M. T. A senescence secretory switch mediated by PI3K/AKT/mTOR activation controls chemoprotective endothelial secretory responses. *Genes Dev.* **30,** 1811–1821 (2016).

209. Höhn, A. *et al.* Happily (n)ever after: Aging in the context of oxidative stress, proteostasis loss and cellular senescence. *Redox Biol* **11,** 482–501 (2016).

210. Lee, B. Y. *et al.* Senescence-associated β-galactosidase is lysosomal β-galactosidase. *Aging Cell* **5,** 187–195 (2006).

211. Debacq-Chainiaux, F., Erusalimsky, J. D., Campisi, J. & Toussaint, O. Protocols to detect senescence-associated beta-galactosidase (SA-βgal) activity, a biomarker of senescent cells in culture and *in vivo*. *Nature Protocols* **4,** 1798–1806 (2009).

212. Childs, B. G., Durik, M., Baker, D. J. & van Deursen, J. M. Cellular senescence in aging and age-related disease: from mechanisms to therapy. *Nat Med* **21,** 1424–1435 (2015).

213. Jun, J.-I. & Lau, L. F. The Matricellular Protein CCN1/CYR61 Induces Fibroblast Senescence and Restricts Fibrosis in Cutaneous Wound Healing. *Nat Cell Biol* **12,** 676–685 (2010).

214. Demaria, M. *et al.* An Essential Role for Senescent Cells in Optimal Wound Healing through Secretion of PDGF-AA. *Dev Cell* **31,** 722–733 (2014).

215. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473,** 337–342 (2011).

216. Protein turnover. - PubMed - NCBI. Available at: https://www.ncbi.nlm.nih.gov/pubmed/16754332/. (Accessed: 1st May 2018)

217. Boisvert, F.-M. *et al.* A Quantitative Spatial Proteomics Analysis of Proteome Turnover in Human Cells. *Mol Cell Proteomics* **11,** (2012).

218. Shen, Y. & Hendershot, L. M. ERdj3, a Stress-inducible Endoplasmic Reticulum DnaJ Homologue, Serves as a CoFactor for BiP's Interactions with Unfolded Substrates. *Mol Biol Cell* **16,** 40–50 (2005).

219. Rhoades, Z. & Ghobarah, H. Broader Coverage and Automatic Mass Calibration Using the TripleTOF™ 5600 System with DuoSpray™ Ion Source. 3

220. Hunter, C. L. & Morrice, N. Microflow SWATH® Acquisition for Industrialized Quantitative Proteomics. 5

221. Yoshida, T., Gan, Q., Shang, Y. & Owens, G. K. Platelet-derived growth factor-BB represses smooth muscle cell marker genes via changes in binding of MKL factors and histone deacetylases to their promoters. *American Journal of Physiology-Cell Physiology* **292,** C886–C895 (2007).

222. Vautrot, V. *et al.* Enhanced SRSF5 Protein Expression Reinforces Lamin A mRNA Production in HeLa Cells and Fibroblasts of Progeria Patients. *Human Mutation* **37,** 280–291 (2016).

223. Arpino, V., Brock, M. & Gill, S. E. The role of TIMPs in regulation of extracellular matrix proteolysis. *Matrix Biology* **44–46,** 247–254 (2015).

224. Karakiulakis, G., Papakonstantinou, E., Aletras, A. J., Tamm, M. & Roth, M. Cell Type-specific Effect of Hypoxia and Platelet-derived Growth Factor-BB on Extracellular Matrix Turnover and Its Consequences for Lung Remodeling. *J. Biol. Chem.* **282,** 908–915 (2007).

225. Circolo, A., Welgus, H. G., Pierce, G. F., Kramer, J. & Strunk, R. C. Differential regulation of the expression of proteinases/antiproteinases in fibroblasts. Effects of interleukin-1 and platelet-derived growth factor. *J. Biol. Chem.* **266,** 12283–12288 (1991).

226. Black, R. A. TIMP3 checks inflammation. *Nature Genetics* **36,** 934–935 (2004).

227. Gardner, J. & Ghorpade, A. Tissue Inhibitor of Metalloproteinase (TIMP)-1: The TIMPed Balance of Matrix Metalloproteinases in the Central Nervous System. *J Neurosci Res* **74,** (2003).

228. Frisch, S. M. & Screaton, R. A. Anoikis mechanisms. *Current Opinion in Cell Biology* **13,** 555–562 (2001).

229. Hou, C.-H., Lin, F.-L., Hou, S.-M. & Liu, J.-F. Cyr61 promotes epithelial-mesenchymal transition and tumor metastasis of osteosarcoma by Raf-1/MEK/ERK/Elk-1/TWIST-1 signaling pathway. *Mol Cancer* **13,** (2014).

230. Chen, X. *et al.* ERK1/2 pathway mediates epithelial-mesenchymal transition by cross-interacting with TGFβ/Smad and Jagged/Notch signaling pathways in lens epithelial cells. *International Journal of Molecular Medicine* **33,** 1664–1670 (2014).

231. Shiozawa, Y. *et al.* GAS6/AXL Axis Regulates Prostate Cancer Invasion, Proliferation, and Survival in the Bone Marrow Niche. *Neoplasia* **12,** 116–127 (2010).

232. Chatzifrangkeskou, M. *et al.* ERK1/2 directly acts on CTGF/CCN2 expression to mediate myocardial fibrosis in cardiomyopathy caused by mutations in the lamin A/C gene. *Hum Mol Genet* **25,** 2220–2233 (2016).

233. Kortlever, R. M., Higgins, P. J. & Bernards, R. Plasminogen activator inhibitor-1 is a critical downstream target of p53 in the induction of replicative senescence. *Nat. Cell Biol.* **8,** 877–884 (2006).

234. Fitzner, B., Lange, A., Müller, S. & Jaster, R. Cdkn1a is a key mediator of rat pancreatic stellate cell senescence. *Pancreatology* **13,** 254–262 (2013).

235. Herbig, U., Wei, W., Dutriaux, A., Jobling, W. A. & Sedivy, J. M. Real-time imaging of transcriptional activation in live cells reveals rapid up-regulation of the cyclin-dependent kinase inhibitor gene CDKN1A in replicative cellular senescence. *Aging Cell* **2,** 295–304 (2003).

236. Siwik, D. A., Pagano, P. J. & Colucci, W. S. Oxidative stress regulates collagen synthesis and matrix metalloproteinase activity in cardiac fibroblasts. *American Journal of Physiology-Cell Physiology* **280,** C53–C60 (2001).

237. Nath, K. A. *et al.* Redox regulation of renal DNA synthesis, transforming growth factor-β1 and collagen gene expression - ScienceDirect. *Kidney International* **53,** 367–381 (1998).

238. Sundaresan, M., Yu, Z.-X., Ferrans, V. J., Irani, K. & Finkel, T. Requirement for Generation of H2O2 for Platelet-Derived Growth Factor Signal Transduction. *Science* **270,** 296–299 (1995).

239. Sun, S.-Y. N-acetylcysteine, reactive oxygen species and beyond. *Cancer Biol Ther* **9,** 109–110 (2010).

240. Kim, K.-Y., Rhim, T., Choi, I. & Kim, S.-S. N-Acetylcysteine Induces Cell Cycle Arrest in Hepatic Stellate Cells through Its Reducing Activity. *J. Biol. Chem.* **276,** 40591–40598 (2001).

241. Zafarullah, M., Li, W. Q., Sylvester, J. & Ahmad, M. Molecular mechanisms of N-acetylcysteine actions. *Cell. Mol. Life Sci.* **60,** 6–20 (2003).

242. Richter, K., Konzack, A., Pihlajaniemi, T., Heljasvaara, R. & Kietzmann, T. Redox-fibrosis: Impact of TGFβ1 on ROS generators, mediators and functional consequences. *Redox Biol* **6,** 344–352 (2015).

243. Conte, E. *et al.* Anti-inflammatory and antifibrotic effects of resveratrol in the lung. *Histol. Histopathol.* **30,** 523–529 (2015).

244. Vrolijk, M. F. *et al.* The shifting perception on antioxidants: the case of vitamin E and β-carotene. *Redox Biol* **4,** 272–278 (2015).

245. Skin cancer statistics. *Cancer Research UK* (2015). Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/skin-cancer. (Accessed: 6th May 2018)

246. Garland, C. F., Garland, F. C. & Gorham, E. D. Epidemiologic Evidence for Different Roles of Ultraviolet A and B Radiation in Melanoma Mortality Rates. *Annals of Epidemiology* **13,** 395–404 (2003).

247. Foote, C. S. DEFINITION OF TYPE I and TYPE II PHOTOSENSITIZED OXIDATION. *Photochemistry and Photobiology* **54,** 659–659 (1991).

248. Sen, C. K., Packer, L. & Baeuerle, P. A. *Antioxidant and Redox Regulation of Genes*. (Academic Press, 1999).

249. Bickers, D. R. & Athar, M. Oxidative Stress in the Pathogenesis of Skin Disease. *Journal of Investigative Dermatology* **126,** 2565–2575 (2006).

250. Wlaschek, M. *et al.* Solar UV irradiation and dermal photoaging. *Journal of Photochemistry and Photobiology B: Biology* **63,** 41–51 (2001).

251. He, Y.-Y., Huang, J.-L. & Chignell, C. F. Delayed and sustained activation of extracellular signal-regulated kinase in human keratinocytes by UVA: implications in carcinogenesis. *J. Biol. Chem.* **279,** 53867–53874 (2004).

252. Yaar, M. & Gilchrest, B. A. Photoageing: mechanism, prevention and therapy. *British Journal of Dermatology* **157,** 874–887 (2007).

253. Baeuerle, P. A. IκB–NF-κB Structures: At the Interface of Inflammation Control. *Cell* **95,** 729–731 (1998).

254. Reelfs, O., Tyrrell, R. M. & Pourzand, C. Ultraviolet A Radiation-Induced Immediate Iron Release Is a Key Modulator of the Activation of NF-κB in Human Skin Fibroblasts. *Journal of Investigative Dermatology* **122,** 1440–1447 (2004).

255. Vile, G. F., Tanew-Iliitschew, A. & Tyrrell, R. M. ACTIVATION OF NF-KB IN HUMAN SKIN FIBROBLASTS BY THE OXIDATIVE STRESS GENERATED BY UVA RADIATION. *Photochemistry and Photobiology* **62,** 463–468 (1995).

256. Pillai, S., Oresajo, C. & Hayward, J. Ultraviolet radiation and skin aging: roles of reactive oxygen species, inflammation and protease activation, and strategies for prevention of inflammation-induced matrix degradation – a review. *International Journal of Cosmetic Science* **27,** 17–34 (2005).

257. Collins, B. C. *et al.* Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nature Communications* **8,** 291 (2017).

258. Challa, A. A. & Stefanovic, B. A Novel Role of Vimentin Filaments: Binding and Stabilization of Collagen mRNAs ▿ . *Mol Cell Biol* **31,** 3773–3789 (2011).

259. Ramsey, C. S. *et al.* Copine-I represses NF-κB transcription by endoproteolysis of p65. *Oncogene* **27,** 3516–3526 (2008).

260. Liu, T., Zhang, L., Joo, D. & Sun, S.-C. NF-κB signaling in inflammation. *Signal Transduction and Targeted Therapy* **2,** 17023 (2017).

261. Fischer, M. Census and evaluation of p53 target genes. *Oncogene* **36,** 3943–3956 (2017).

262. Latonen, L., Taya, Y. & Laiho, M. UV-radiation induces dose-dependent regulation of p53 response and modulates p53-HDM2 interaction in human fibroblasts. *Oncogene* **20,** 6784–6793 (2001).

263. Farrukh, M. R. *et al.* Oxidative stress mediated Ca(2+) release manifests endoplasmic reticulum stress leading to unfolded protein response in UV-B irradiated human skin cells. *J. Dermatol. Sci.* **75,** 24–35 (2014).

264. Zamansky, G. B. & Chou, I.-N. Environmental Wavelengths of Ultraviolet Light Induce Cytoskeletal Damage. *Journal of Investigative Dermatology* **89,** 603–606 (1987).

265. Veselska, R. & Janisch, R. The Effect of UV Radiation on Changes in Cytoskeleton and Viability of Mouse Fibroblasts L929 Cell Line. *Scripta Medica (BRNO)* **73,** 393–408 (2000).

266. Szabo, R., Netzel-Arnett, S., Hobson, J. P., Antalis, T. M. & Bugge, T. H. Matriptase-3 is a novel phylogenetically preserved membrane-anchored serine protease with broad serpin reactivity. *Biochem J* **390,** 231–242 (2005).

267. Jin, X. *et al.* Matriptase activates stromelysin (MMP-3) and promotes tumor growth and angiogenesis. *Cancer Science* **97,** 1327–1334 (2006).

268. Decorsière, A., Cayrel, A., Vagner, S. & Millevoi, S. Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage. *Genes Dev.* **25,** 220–225 (2011).

269. Dreyfuss, G., Matunis, M. J., Pinol-Roma, S. & Burd, C. G. hnRNP Proteins and the Biogenesis of mRNA. *Annual Review of Biochemistry* **62,** 289–321 (1993).

270. Uren, P. J. *et al.* High-throughput analyses of hnRNP H1 dissects its multi-functional aspect. *RNA Biol* **13,** 400–411 (2016).

271. Gautrey, H. *et al.* SRSF3 and hnRNP H1 regulate a splicing hotspot of HER2 in breast cancer cells. *RNA Biol* **12,** 1139–1151 (2015).

272. Bost, F., Diarra-Mehrpour, M. & Martin, J.-P. Inter-α-trypsin inhibitor proteoglycan family. *European Journal of Biochemistry* **252,** 339–346 (1998).

273. Paris, S. *et al.* Inhibition of tumor growth and metastatic spreading by overexpression of inter-alpha-trypsin inhibitor family chains. *International Journal of Cancer* **97,** 615–620 (2002).

274. Thurstan, S. A. *et al.* Chemical consequences of cutaneous photoageing. *Chem Cent J* **6,** 34 (2012).

275. Stapelberg, M. P. F., Williams, R. B. H., Byrne, S. N. & Halliday, G. M. The Alternative Complement Pathway Seems to be a UVA Sensor that Leads to Systemic Immunosuppression. *Journal of Investigative Dermatology* **129,** 2694–2701 (2009).

276. Tomkowicz, B. *et al.* Interaction of endosialin/TEM1 with extracellular matrix proteins mediates cell adhesion and migration. *PNAS* **104,** 17965–17970 (2007).

277. Chen, N., Chen, C.-C. & Lau, L. F. Adhesion of Human Skin Fibroblasts to Cyr61 Is Mediated through Integrin α6β1 and Cell Surface Heparan Sulfate Proteoglycans. *J. Biol. Chem.* **275,** 24953–24961 (2000).

278. Chen, C.-C., Chen, N. & Lau, L. F. The Angiogenic Factors Cyr61 and Connective Tissue Growth Factor Induce Adhesive Signaling in Primary Human Skin Fibroblasts. *J. Biol. Chem.* **276,** 10443–10452 (2001).

279. Todorovicç, V., Chen, C.-C., Hay, N. & Lau, L. F. The matrix protein CCN1 (CYR61) induces apoptosis in fibroblasts. *J Cell Biol* **171,** 559–568 (2005).

280. Shimoda, M. *et al.* Hyaluronan-Binding Protein Involved in Hyaluronan Depolymerization Controls Endochondral Ossification through Hyaluronan Metabolism. *Am. J. Pathol.* **187,** 1162–1176 (2017).

281. Aya, K. L. & Stern, R. Hyaluronan in wound healing: Rediscovering a major player. *Wound Repair and Regeneration* **22,** 579–593 (2014).

282. Watson, R. E. B., Gibbs, N. K., Griffiths, C. E. M. & Sherratt, M. J. Damage to Skin Extracellular Matrix Induced by UV Exposure. *Antioxidants & Redox Signaling* **21,** 1063–1077 (2013).

283. Herrmann, G. *et al.* UVA irradiation stimulates the synthesis of various matrix-metalloproteinases (MMPs) in cultured human fibroblasts. *Experimental Dermatology* **2,** 92–97 (1993).

284. Murata, H. *et al.* Glutaredoxin Exerts an Antiapoptotic Effect by Regulating the Redox State of Akt. *J. Biol. Chem.* **278,** 50226–50233 (2003).

285. Huang, X. *et al.* Crystal Structure of an Inactive Akt2 Kinase Domain. *Structure* **11,** 21–30 (2003).

286. Jin, C.-Y. *et al.* Sulforaphane sensitizes tumor necrosis factor-related apoptosis-inducing ligand-mediated apoptosis through downregulation of ERK and Akt in lung adenocarcinoma A549 cells. *Carcinogenesis* **28,** 1058–1066 (2007).

287. Favaro, E. *et al.* MicroRNA-210 Regulates Mitochondrial Free Radical Response to Hypoxia and Krebs Cycle in Cancer Cells by Targeting Iron Sulfur Cluster Protein ISCU. *PLOS ONE* **5,** e10345 (2010).

288. Salajegheh, A. Iron-Sulfur Clusters (ISCU). in *Angiogenesis in Health, Disease and Malignancy* (Springer, 2016).

289. Reunanen, N., Foschi, M., Han, J. & Kähäri, V.-M. Activation of Extracellular Signal-regulated Kinase 1/2 Inhibits Type I Collagen Expression by Human Skin Fibroblasts. *J. Biol. Chem.* **275,** 34634–34639 (2000).

290. Kim, H. J. *et al.* IL-18 Downregulates Collagen Production in Human Dermal Fibroblasts via the ERK Pathway. *Journal of Investigative Dermatology* **130,** 706–715 (2010).

291. Papaioannou, I. *et al.* STAT3 controls COL1A2 enhancer activation cooperatively with JunB, regulates type I collagen synthesis posttranscriptionally, and is essential for lung myofibroblast differentiation. *MBoC* **29,** 84–95 (2018).

292. Chung, J., Uchida, E., Grammer, T. C. & Blenis, J. STAT3 serine phosphorylation by ERK-dependent and -independent pathways negatively modulates its tyrosine phosphorylation. *Mol. Cell. Biol.* **17,** 6508–6516 (1997).

293. Ley, R., Balmanno, K., Hadfield, K., Weston, C. & Cook, S. J. Activation of the ERK1/2 Signaling Pathway Promotes Phosphorylation and Proteasome-dependent Degradation of the BH3-only Protein, Bim. *J. Biol. Chem.* **278,** 18811–18816 (2003).

294. Nimni, M. E. Fibrillar Collagens: Their Biosynthesis, Molecular Structure, and Mode of Assembly. in *Extracellular Matrix: Chemistry: Biology, and Pathobiology with Emphasis on the Liver* 121–149 (CRC Press, 1993).

295. Lodish, H. *et al.* Collagen: The Fibrous Proteins of the Matrix. in *Molecular Cell Biology* (W. H. Freeman, 2000).

296. Chen, L., Mao, S. J. & Larsen, W. J. Identification of a factor in fetal bovine serum that stabilizes the cumulus extracellular matrix. A role for a member of the inter-alpha-trypsin inhibitor family. *J. Biol. Chem.* **267,** 12380–12386 (1992).

297. Sandson, J., Hamerman, D. & Schwick, G. Altered properties of pathological hyaluronate due to a bound inter-alpha trypsin inhibitor. *Trans. Assoc. Am. Physicians* **78,** 304–313 (1965).

298. HUTADILOK, N., GHOSH, P. & BROOKS, P. M. Binding of haptoglobin, inter-a-trypsin inhibitor, and a, proteinase inhibitor to synovial fluid hyaluronate and the influence of these proteins on its degradation by oxygen derived free radicals. 10

299. Onodera, Y., Teramura, T., Takehara, T. & Fukuda, K. Hyaluronic acid regulates a key redox control factor Nrf2 via phosphorylation of Akt in bovine articular chondrocytes. *FEBS Open Bio* **5,** 476–484 (2015).

300. Clydesdale, G. J., Dandie, G. W. & Muller, H. K. Ultraviolet light induced injury: Immunological and inflammatory effects. *Immunology and Cell Biology* **79,** 547–568 (2001).

301. Zoltan-Jones, A., Huang, L., Ghatak, S. & Toole, B. P. Elevated Hyaluronan Production Induces Mesenchymal and Transformed Properties in Epithelial Cells. *J. Biol. Chem.* **278,** 45801–45810 (2003).

302. Jenkins, R. H., Thomas, G. J., Williams, J. D. & Steadman, R. Myofibroblastic Differentiation Leads to Hyaluronan Accumulation through Reduced Hyaluronan Turnover. *J. Biol. Chem.* **279,** 41453–41460 (2004).

303. Darby, I. A., Laverdet, B., Bonté, F. & Desmoulière, A. Fibroblasts and myofibroblasts in wound healing. *Clin Cosmet Investig Dermatol* **7,** 301–311 (2014).

304. Stevens, K. L. P. *et al.* Diminished Ost3-dependent N-glycosylation of the BiP nucleotide exchange factor Sil1 is an adaptive response to reductive ER stress. *Proceedings of the National Academy of Sciences* **114,** 12489–12494 (2017).

305. Liu, Y. *et al.* Quantitative measurements of N-linked glycoproteins in human plasma by SWATH-MS. *PROTEOMICS* **13,** 1247–1256 (2013).

306. Zhang, Y. *et al.* The Use of Variable Q1 Isolation Windows Improves Selectivity in LC–SWATH–MS Acquisition. *J. Proteome Res.* **14,** 4359–4371 (2015).

307. What is Variable Window Acquisition and why does it improve results quality? Available at: https://sciex.com/community/application-discussions/proteomics/swath/data-acquisition/what-is-variable-window-acquisition-and-why-does-it-improve-results-quality. (Accessed: 10th May 2018)

308. Kuhn, E. *et al.* Developing Multiplexed Assays for Troponin I and Interleukin-33 in Plasma by Peptide Immunoaffinity Enrichment and Targeted Mass Spectrometry. *Clinical Chemistry* **55,** 1108–1117 (2009).

309. Hoofnagle, A. N., Becker, J. O., Wener, M. H. & Heinecke, J. W. Quantification of Thyroglobulin, a Low-abundance Serum Protein, by Immunoaffinity Peptide Enrichment and Tandem Mass Spectrometry. *Clin Chem* **54,** 1796–1804 (2008).

310. Schubert, O. T. *et al.* Absolute Proteome Composition and Dynamics during Dormancy and Resuscitation of Mycobacterium tuberculosis. *Cell Host & Microbe* **18,** 96–108 (2015).

311. Ludwig, C., Claassen, M., Schmidt, A. & Aebersold, R. Estimation of Absolute Protein Quantities of Unlabeled Samples by Selected Reaction Monitoring Mass Spectrometry. *Mol Cell Proteomics* **11,** M111.013987 (2012).

312. Batth, T. S. *et al.* Large-Scale Phosphoproteomics Reveals Shp-2 Phosphatase-Dependent Regulators of Pdgf Receptor Signaling. *Cell Reports* **22,** 2784–2796 (2018).

313. Ghezzi, P. & Chan, P. Redox Proteomics Applied to the Thiol Secretome. *Antioxidants & Redox Signaling* **26,** 299–312 (2017).

314. Boronat, S., Domènech, A. & Hidalgo, E. Proteomic Characterization of Reversible Thiol Oxidations in Proteomes and Proteins. *Antioxidants & Redox Signaling* **26,** 329–344 (2017).

Appendix 1: R scripts

The scripts presented here were written by the author of this thesis.

Contents

GO analysis

#set working directory to local folder
```
setwd("C://Users/naomi/Documents/R")
library(clusterProfiler)
library(org.Hs.eg.db)
```

#load clusterprofiler and human annotations, then create vector with IDs required as GeneName from MS Excel export
```
GeneNameTest <- read.csv("~/R/GeneNameTest.csv")
View(GeneNameTest)
```

#Create vector with just Gene Name as IDs
```
IDs <- as.vector(GeneNameTest$Gene.Name)
```

#convert IDs from gene name to entrezid
```
ConvertedIDs = bitr(IDs, fromType = "SYMBOL", toType = "ENTREZID", OrgDb
= "org.Hs.eg.db")
```

#Identify unmapped gene names
```
lost <- GeneNameTest$Gene.Name[!(GeneNameTest$Gene.Name %in%
ConvertedIDs$SYMBOL)]
```

#Correct unmapped gene names in excel file
```
GeneNameTestCorrected <- read.csv("~/R/GeneNameTestCorrected.csv")
```

#Re-run vector formation and convertion to entrezids
```
CorrectedIDs <- as.vector(GeneNameTestCorrected$Corrected.Gene.Name)
ConvertedCorrectedIDs =bitr(CorrectedIDs, fromType = "SYMBOL", toType =
"ENTREZID", OrgDb = "org.Hs.eg.db")
```

#Keep checking until 1:1 mapping returned. This may require searching for Entrez ID online
#Make sure when looking for lost again that vector names etc are updated
#Make vector of corrected and converted IDs
```
Genes <- c(ConvertedCorrectedIDs[[2]])
```

#Use GroupGO to annotate. BP = biological process, CC = cellular `compartment`, `MF = molecular function`

#Use ShowCategory to change how many bars on the graph
```
GenesGOCC <- groupGO(Genes, OrgDb = "org.Hs.eg.db", keytype =
"ENTREZID", ont = "CC", level = 2,readable = FALSE)
barplot(GenesGOCC, drop=TRUE, showCategory=12)
GenesGOBP <-groupGO(Genes, OrgDb = "org.Hs.eg.db", keytype = "ENTREZID",
ont = "BP", level = 2,readable = FALSE)
barplot(GenesGOBP, drop=TRUE, showCategory=12)
```

Box plots

```
#Set working directory and load required packages
setwd("C:/Users/naomi/Documents/Postgrad/Thesis MS Data")
library(ggplot2)
library(reshape2)

#Nano/micro comparison
NanoMicro <- read.csv("NanoMicro.csv")
boxplot(NanoMicro, range = 0, ylab= "Number of peptides", xlab= "Flow source")


#replicates analysis PDGFDTT to PDGF only
PDGFDTTPDGFpeaksall <- read.csv("PDGFDTTPDGF peaksall.csv")
PDGFDTTPDGFpeaksallBP <- melt(PDGFDTTPDGFpeaksall, id.vars = "Gene.Name")
PDGFDTTBP <- PDGFDTTPDGFpeaksallBP[c(1:252),]
PDGFDTTBP$variable="PDGF and DTT"
PDGFonlyBP <- PDGFDTTPDGFpeaksallBP[c(253:504),]
PDGFonlyBP$variable="PDGF only"
PDGFDTTBPtotal <- rbind(PDGFDTTBP,PDGFonlyBP)


ggplot(PDGFDTTBPtotal, aes(x=Gene.Name, y=value, fill=variable)) +
  geom_boxplot() + labs (x=" Gene Name", y= "Peak area")+ theme_classic()+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())


#PDGFDTT technical replicates box plot
PDGFDTTPDGFpeaksreps <- read.csv("PDGFDTTPDGF peaksreplicates.csv")
PDGFDTTPDGFpeaksrepsBP <- melt(PDGFDTTPDGFpeaksreps, id.vars = "Gene.Name")
PDGFDTT1BP <- PDGFDTTPDGFpeaksrepsBP[c(1:84),]
PDGFDTT1BP$variable="Biological replicate 1"
PDGFDTT2BP <- PDGFDTTPDGFpeaksrepsBP[c(85:168),]
PDGFDTT2BP$variable="Biological replicate 2"
PDGFDTT3BP <- PDGFDTTPDGFpeaksrepsBP[c(169:252),]
PDGFDTT3BP$variable="Biological replicate 3"
PDGFDTTrepsBPtotal <- rbind(PDGFDTT1BP, PDGFDTT2BP, PDGFDTT3BP)

ggplot(PDGFDTTrepsBPtotal, aes(x=Gene.Name, y=value, fill=variable)) +
  geom_boxplot() + labs (x=" Gene Name", y= "Peak area")+ theme_classic()+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), text =
element_text(size=20),
        axis.text.x = element_text(angle=90, hjust=1)) + theme(legend.justification = c(1, 1),
legend.position = c(1, 1))
```

```
#PDGF only technical replicates box plot
PDGFonly1BP <- PDGFDTTPDGFpeaksrepsBP[c(253:336),]
PDGFonly1BP$variable="Biological replicate 1"
PDGFonly2BP <- PDGFDTTPDGFpeaksrepsBP[c(337:420),]
PDGFonly2BP$variable="Biological replicate 2"
PDGFonly3BP <- PDGFDTTPDGFpeaksrepsBP[c(421:504),]
PDGFonly3BP$variable="Biological replicate 3"
PDGFonlyrepsBPtotal <- rbind(PDGFonly1BP, PDGFonly2BP, PDGFonly3BP)


ggplot(PDGFonlyrepsBPtotal, aes(x=Gene.Name, y=value, fill=variable)) +
  geom_boxplot() + labs (x=" Gene Name", y= "Peak area")+ theme_classic()+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), text =
element_text(size=20),
      axis.text.x = element_text(angle=90, hjust=1)) + theme(legend.justification = c(1, 1),
legend.position = c(1, 1))


#replicates analysis DTT to Untreated box plot
DTTUntreatedpeaksall <- read.csv("DTTUntreated peaksall.csv")
DTTUntreatedpeaksallBP <- melt(DTTUntreatedpeaksall, id.vars = "Gene.Name")
DTTBP <- DTTUntreatedpeaksallBP[c(1:405),]
DTTBP$variable="DTT"
UntreatedBP <- DTTUntreatedpeaksallBP[c(406:810),]
UntreatedBP$variable="Untreated"
DTTUntreatedBPtotal <- rbind(DTTBP,UntreatedBP)

ggplot(DTTUntreatedBPtotal, aes(x=Gene.Name, y=value, fill=variable)) +
  geom_boxplot() + labs (x=" Gene Name", y= "Peak area")+ theme_classic()+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank())

#DTT technical replicates box plot
DTTUpeaksreps <- read.csv("DTTUntreated peaksreplicates.csv")
DTTUpeaksrepsBP <- melt(DTTUpeaksreps, id.vars = "Gene.Name")
DTT1BP <- DTTUpeaksrepsBP[c(1:135),]
DTT1BP$variable="Biological replicate 1"
DTT2BP <- DTTUpeaksrepsBP[c(136:270),]
DTT2BP$variable="Biological replicate 2"
DTT3BP <- DTTUpeaksrepsBP[c(271:405),]
DTT3BP$variable="Biological replicate 3"
DTTrepsBPtotal <- rbind(DTT1BP, DTT2BP, DTT3BP)

ggplot(DTTrepsBPtotal, aes(x=Gene.Name, y=value, fill=variable)) +
  geom_boxplot() + labs (x=" Gene Name", y= "Peak area")+ theme_classic()+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), text =
element_text(size=20),
      axis.text.x = element_text(angle=90, hjust=1)) + theme(legend.justification = c(1, 1),
legend.position = c(1, 1))
```

```
#Untreated technical replicates box plot
Untreated1BP <- DTTUpeaksrepsBP[c(406:540),]
Untreated1BP$variable="Biological replicate 1"
Untreated2BP <- DTTUpeaksrepsBP[c(541:675),]
Untreated2BP$variable="Biological replicate 2"
Untreated3BP <- DTTUpeaksrepsBP[c(676:810),]
Untreated3BP$variable="Biological replicate 3"
UntreatedrepsBPtotal <- rbind(Untreated1BP, Untreated2BP, Untreated3BP)

ggplot(UntreatedrepsBPtotal, aes(x=Gene.Name, y=value, fill=variable)) +
  geom_boxplot() + labs (x=" Gene Name", y= "Peak area")+ theme_classic()+
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(), text =
element_text(size=20),
      axis.text.x = element_text(angle=90, hjust=1)) + theme(legend.justification = c(1, 1),
legend.position = c(1, 1))
```

Venn diagrams

Pairwise Venn – Nano to Micro flow comparison

```
setwd("C:/Users/naomi/Documents/Postgrad/Thesis MS Data")
NanoMicrolists <- read.csv("NanoMicrolists.csv")
Nano <- c(as.character(NanoMicrolists$Nano))
Micro <- c(as.character(NanoMicrolists$Micro))
NanonotMicro <- c(setdiff(Nano, Micro))
MicronotNano <- c(setdiff(Micro, Nano))
library(VennDiagram)
draw.pairwise.venn(4487,4982,3759, category=c("Nano flow", "Micro
flow"), scaled = FALSE, ext.text = FALSE, cex=c(2,2,2), fontface =
rep("bold", 3), cat.cex=c(1.5,1.5), cat.fontface = rep("bold",2),
cat.pos=c(315,45), cat.dist=c(0.13,0.13))
```

Quadruple Venn

```
library(VennDiagram)
library(gridExtra)
library(ggplot2)

setwd("C:/Users/naomi/Documents/R")
PComb <- read.csv("PComb.csv")
PCombC <- c(as.character(PComb$P))
PDComb <- read.csv("PDComb.csv")
PDCombC <- c(as.character(PDComb$PD))
DComb <- read.csv("DComb.csv")
DCombC <- c(as.character(DComb$D))
UComb <- read.csv("UComb.csv")
UCombC <- c(as.character(UComb$U))
OverlapComb <- calculate.overlap(x=list("PDCombC" = PDCombC, "PCombC" =
PCombC, "DCombC" = DCombC, "UCombC" = UCombC))
#draw quad venn. Using numbers: (area1, area2, area3, area4,
(SUMa2,11,12,6), (SUMa4,5,6,12), (SUMa5,6,10,11), (SUMa6,7,12,13),
(SUMa6,7,8,11), (SUMa5,6,7,15),(SUMa12,6),(SUMa11,6), (a5+a6), (a7+a6),
a6)

gR1 = draw.quad.venn(2060, 1996, 2035,
2044,(43+51+57+1629),(52+89+1629+57),(89+1629+41+51), (1629+52+57+28),
(1629+52+48+51), (89+1629+52+38), (1629+57), (51+1629), (1629+89),
(52+1629), 1629, category = c("PDGF and DTT", "PDGF only", "DTT only",
"Untreated"), col = "black", fill = c("#DF3034","#28A197","#FFBF47",
"#912B88"),fontface = rep("bold", 15),fontfamily = rep("sans", 15), cex
= rep(2,15), cat.cex = rep(2, 4), cat.fontface = rep("bold", 4),
cat.fontfamily = rep("sans", 4))
require(gridExtra)
grid.arrange(gTree(children=gR1), top=textGrob("Combined Repeats",
gp=gpar(fontsize=24, fontface = "bold", fontfamily = "sans")))
```

P-value adjustment

```
p.values <- read.csv(“.csv”)
p.values$p.adjusted <- (p.adjust(p.values$p, method = “fdr”)
write.csv(p.values, “adjustedpvalues.csv”)
```

#Where p.values is dual column csv of gene name and corresponding p values. FC can also be included but are irrelevant in this script. “.csv” should be replaced with desired file name.

Volcano plots

Colour coded volcano plots

```
#set working directory and load required packages
setwd("C:/Users/naomi/Documents/Postgrad/Thesis MS Data/UVA")

library(org.Hs.eg.db)
library(clusterProfiler)
library(ggplot2)
library(dplyr)
library(ggrepel)

#import data from file and transform relevant vectors as required
UVA <- read.csv("UVCVFP.csv")
UVA$neglog.p.adj = -log10(UVA$p.adjusted)

#subset data
UVAsig <- subset(UVA,UVA$neglog.p.adj>1.3 &
UVA$iLog.FC>0.3|UVA$iLog.FC<(-0.3))
UVAsigsig <- subset(UVAsig, UVAsig$p.adjusted<0.05)
UVAsigsigCVC <- subset(UVAsigsig, UVAsigsig$CVC<30)
UVAsigsigCVCCVUV <- subset(UVAsigsigCVC, UVAsigsigCVC$CVUV<30)

#create volcano plot data frame
VolcanoUVA = mutate(UVA, sig=ifelse(UVA$neglog.p.adj>1.3 &
(UVA$iLog.FC>0.3|UVA$iLog.FC<(-0.3)), "Significant", "Not"))
Volcanovariance = mutate(VolcanoUVA,
variance=ifelse(VolcanoUVA$CVC<30&(VolcanoUVA$CVUV<30),
"tight","variable"))
VolcanoUVAcoloured = mutate(Volcanovariance,
key=ifelse(grepl("Significant",sig)&(grepl("tight", variance)),"%CV<30,
p<0.05 and fold change>2",
ifelse(grepl("Significant",sig)&(grepl("variable", variance)), "p<0.05
and fold change>2", "not significant")))

#create volcano plot
ggplot(VolcanoUVAcoloured,aes(iLog.FC, (neglog.p.adj))) +
  geom_point(aes(col=key), size=2) +
  scale_color_manual(name = "Significance filter", values=c("#7E317B",
"dark grey", "#D8ACE0"), breaks=c("%CV<30, p<0.05 and fold change>2",
"p<0.05 and fold change>2", "not significant")) +
  theme_classic() +
  xlab("Log(10) Fold Change") +
  ylab("-log(10) Adjusted p-value") +
  ggtitle("UV to control comparison")+
  theme(plot.title = element_text(hjust = 0.5))+
  theme(legend.title=element_blank())+
  theme(text = element_text(size=20)) +
  theme(axis.line = element_line(size=1.25)) +
  theme(plot.margin=unit(c(0.5,0.5,1,0.5),"cm"))
```

Significant collagen labelled volcano plots

#set working directory and install relevant packages

```
setwd("C:/Users/naomi/Documents/Postgrad/Thesis MS Data/PDGFDTT Micro")
library(org.Hs.eg.db)
library(clusterProfiler)
library(ggplot2)
library(dplyr)
library(ggrepel)
```

#DTT to Untreated comparison

#import data from file and convert to numerical vectors

```
DTTUntreated <- read.csv("DTTUntreatedComparison.csv")
DTTUntreated$p.valueNumber =
as.numeric(as.character(DTTUntreated$p.value.adjusted))
DTTUntreated$logp.valueNumber = -log10(DTTUntreated$p.valueNumber)
DTTUntreated$Fold.ChangeNumber =
as.numeric(as.character(DTTUntreated$Fold.Change))
DTTUntreated$Log.Fold.ChangeNumber =
log10(DTTUntreated$Fold.ChangeNumber)
```

#assign labels to collagen genes and all others

```
DTTUntreated$group <- "not important"
DTTUntreated$group[DTTUntreated$Gene.Name.1 %in%
c("COL1A1","COL2A1","COL3A1", "COL5A1", "COL6A1", "COL7A1", "COL8A1",
"COL11A1", "COL12A1", "COL14A1", "COL15A1", "COL18A1", "COL1A2",
"COL4A2", "COL5A2", "COL6A2", "TAGLN", "COL6A3")] <- "Collagen"
```

#subset data

```
DTTUntreatedsub <- subset(DTTUntreated, DTTUntreated$group=="Collagen")
DTTUntreatedsub1.5 <- subset(DTTUntreatedsub,
DTTUntreatedsub$logp.valueNumber>1.3 &
DTTUntreatedsub$Log.Fold.ChangeNumber<(-0.176)&
DTTUntreatedsub$Log.Fold.ChangeNumber>(-0.30))

DTTUntreatedsub2 <- subset(DTTUntreatedsub,
DTTUntreatedsub$logp.valueNumber>1.3 &
DTTUntreatedsub$Log.Fold.ChangeNumber>0.3|DTTUntreatedsub$Log.Fold.Chang
eNumber<(-0.3))
```

#create data frame for volcano plot

```
VolcanoDTTUntreatedlogs = mutate(DTTUntreated,
sig=ifelse(DTTUntreated$logp.valueNumber>1.3 &
(DTTUntreated$Log.Fold.ChangeNumber>0.3|DTTUntreated$Log.Fold.ChangeNumb
er<(-0.3)), "P<0.05 and Fold Change>2", "Not Significant"))
```

```
#create volcano plot
vDTTUlogs = ggplot(VolcanoDTTUntreatedlogs,
                   aes(Log.Fold.ChangeNumber, (logp.valueNumber))) +
  geom_point(aes(col=sig), size=2) +
  scale_color_manual(values=c("grey", "black")) +
  theme_classic() +
  xlab("Log(10) Fold Change") +
  ylab("-log(10) Adjusted p-value") +
  ggtitle("DTT to Untreated Comparison")+
  theme(plot.title = element_text(hjust = 0.5))+
  theme(legend.title=element_blank())+
  theme(text = element_text(size=20)) +
  theme(axis.line = element_line(size=1.25)) +
  theme(plot.margin=unit(c(0.5,0.5,1,0.5),"cm")) +
  theme(legend.justification = c(1,1), legend.position = c(1,0.75)) +
  xlim(-1,1.5) +
  geom_text_repel(data = DTTUntreatedsub2, aes(label = Gene.Name.1),
hjust = 0, size = 4, min.segment.length = 0, fontface = "bold") +
  geom_text_repel(data = DTTUntreatedsub1.5, aes(label = Gene.Name.1),
hjust = 0, size = 4, min.segment.length = 0)

vDTTUlogs


#PDGF and DTT to PDGF only comparison
#import data from file and convert to numerical vectors
PDGFDTTPDGF <- read.csv("PDGFDTTDTTComparison.csv")
PDGFDTTPDGF$p.valueNumber =
as.numeric(as.character(PDGFDTTPDGF$p.value.adjusted))
PDGFDTTPDGF$logp.valueNumber = -log10(PDGFDTTPDGF$p.valueNumber)
PDGFDTTPDGF$Fold.ChangeNumber =
as.numeric(as.character(PDGFDTTPDGF$Fold.Change))
PDGFDTTPDGF$Log.Fold.ChangeNumber = log10(PDGFDTTPDGF$Fold.ChangeNumber)

#assign labels to collagen genes and all others
PDGFDTTPDGF$group <- "not important"
PDGFDTTPDGF$group[PDGFDTTPDGF$Gene.Name.1 %in%
c("COL1A1","COL2A1","COL3A1", "COL5A1", "COL6A1", "COL7A1", "COL8A1",
"COL11A1", "COL12A1", "COL14A1", "COL15A1", "COL18A1", "COL1A2",
"COL4A2", "COL5A2", "COL6A2", "TAGLN", "COL6A3")] <- "Collagen"

#subset significant collagens as those with p<0.05 and FC>1.5 or FC>2
PDGFDTTPDGFsub <- subset(PDGFDTTPDGF, PDGFDTTPDGF$group=="Collagen")
PDGFDTTPDGFsub1.5 <- subset(PDGFDTTPDGFsub,
PDGFDTTPDGFsub$logp.valueNumber>1.3 &
PDGFDTTPDGFsub$Log.Fold.ChangeNumber<(-0.176)&
PDGFDTTPDGFsub$Log.Fold.ChangeNumber>(-0.30))
PDGFDTTPDGFsub2 <- subset(PDGFDTTPDGFsub,
PDGFDTTPDGFsub$logp.valueNumber>1.3 &
PDGFDTTPDGFsub$Log.Fold.ChangeNumber>0.3|PDGFDTTPDGFsub$Log.Fold.ChangeN
umber<(-0.3))

#create data frame for volcano plot
VolcanoPDGFDTTPDGFlogs = mutate(PDGFDTTPDGF,
sig=ifelse(PDGFDTTPDGF$logp.valueNumber>1.3 &
```

```
(PDGFDTTPDGF$Log.Fold.ChangeNumber>0.3|PDGFDTTPDGF$Log.Fold.ChangeNumber
<(-0.3)), "P<0.05 and Fold Change>2", "Not Significant"))

#create volcano plot
vPDGFDTTPDGFlogs = ggplot(VolcanoPDGFDTTPDGFlogs,
                          aes(Log.Fold.ChangeNumber, (logp.valueNumber)))
+
  geom_point(aes(col=sig), size=2) +
  scale_color_manual(values=c("grey", "black")) +
  theme_classic() +
  xlab("Log(10) Fold Change") +
  ylab("-log(10) Adjusted p-value") +
  ggtitle("PDGF and DTT to PDGF only Comparison")+
  theme(plot.title = element_text(hjust = 0.5))+
  theme(legend.title=element_blank())+
  theme(text = element_text(size=20)) +
  theme(axis.line = element_line(size=1.25))+
  theme(plot.margin=unit(c(0.5,0.5,1,0.5),"cm"))+
  theme(legend.justification = c(1,1), legend.position = c(1,0.75)) +
  xlim(-1,1.5) +
  geom_text_repel(data = PDGFDTTPDGFsub2, aes(label = Gene.Name.1),
hjust = 0, size = 4, min.segment.length = 0, fontface = "bold") +
  geom_text_repel(data = PDGFDTTPDGFsub1.5, aes(label = Gene.Name.1),
hjust = 0, size = 4, min.segment.length = 0)

vPDGFDTTPDGFlogs
```

Heatmap

Comparison of significantly changing proteins

```
setwd("C:/Users/naomi/Documents/Postgrad/Thesis MS Data/PDGFDTT Micro")
library(ggplot2)
library(reshape2)

SignificantDTTU <- read.csv("SignificantDTTU.csv")
names(SignificantDTTU)[1:3] <- c("Gene.Name", "DTT/Untreated",
"PDGFDTT/PDGFonly")

SignificantPDGFDTTPDGF <- read.csv("SignificantPDGFDTTPDGF.csv")
names(SignificantPDGFDTTPDGF)[1:2] <- c("Gene.Name", "PDGFDTT/PDGFonly")

heatmapDF <- melt(SignificantDTTU[1:2], id.vars = "Gene.Name")
names(heatmapDF)[2:3] <- c("Comparison", "Fold_change")

heatmapDFPDGF <- melt(SignificantPDGFDTTPDGF, id.vars = "Gene.Name")
names(heatmapDFPDGF)[2:3] <- c("Comparison", "Fold_change")

ggplot(heatmapDF, aes(x=Comparison, y=reorder(Gene.Name, Fold_change,
order=TRUE))) +
  geom_tile(aes(fill = Fold_change), color = "white") +
  scale_fill_gradient(low = "white", high = "#7E317B", na.value =
"black") +
  ylab("Gene Name") +
  xlab("Comparison") +
  theme(legend.title = element_text(size = 10),
        legend.text = element_text(size = 12),
        plot.title = element_text(size=16),
        axis.title=element_text(size=14,face="bold"),
        axis.text.x = element_text(angle = 90, hjust = 1)) +
  labs(fill = "Fold Change") + theme_minimal() +
  theme(panel.grid.major = element_blank(), panel.grid.minor =
element_blank()) +
  scale_x_discrete(expand = c(0, 0)) + scale_y_discrete(expand = c(0,
0)) +
  theme(axis.ticks = element_blank())

ggplot(heatmapDFPDGF, aes(x=Comparison, y=reorder(Gene.Name,
Fold_change, order=TRUE))) +
  geom_tile(aes(fill = Fold_change), color = "white") +
  scale_fill_gradient(low = "white", high = "#7E317B", na.value =
"black") +
  ylab("Gene Name") +
  xlab("Comparison") +
  theme(legend.title = element_text(size = 10),
        legend.text = element_text(size = 12),
        plot.title = element_text(size=16),
        axis.title=element_text(size=14,face="bold"),
        axis.text.x = element_text(angle = 90, hjust = 1)) +
  labs(fill = "Fold Change") + theme_minimal() +
  theme(panel.grid.major = element_blank(), panel.grid.minor =
element_blank()) +
  scale_x_discrete(expand = c(0, 0)) + scale_y_discrete(expand = c(0,
0)) +
  theme(axis.ticks = element_blank())
```

Comparison of collagen proteins

```
setwd("C:/Users/naomi/Documents/Postgrad/Thesis MS Data/UVA")
library(ggplot2)
library(reshape2)

Collagens <- read.csv("iCollagen comparison.csv")
names(Collagens)[1:3] <- c("Gene.Name", "DTT/Untreated",
"UVA/Untreated")
heatmapCollagens <- melt(Collagens, id.vars = "Gene.Name")
heatmapCollagens$forder <- as.numeric(1:17)


ggplot(heatmapCollagens, aes(x=variable, y=reorder(Gene.Name, forder,
median, order=TRUE))) +
  geom_tile(aes(fill = value), color = "white") +
  scale_fill_gradient2(low = ("#006388"), mid = "white",
                       high = ("#7E317B"), midpoint = 1, space = "Lab",
                       na.value = "grey50", guide = "colourbar",
aesthetics = "fill") +
  ylab("Gene Name") +
  xlab("Comparison") +
  theme(legend.title = element_text(size = 10),
        legend.text = element_text(size = 12),
        plot.title = element_text(size=16),
        axis.title=element_text(size=14,face="bold"),
        axis.text.x = element_text(angle = 90, hjust = 1)) +
  labs(fill = "Fold Change") + theme_minimal() +
  theme(panel.grid.major = element_blank(), panel.grid.minor =
element_blank()) +
  scale_x_discrete(expand = c(0, 0)) + scale_y_discrete(expand = c(0,
0)) +
  theme(axis.ticks = element_blank())
```